# TEACHERS' PERCEPTIONS OF THE GOAL-SETTING MODEL AND THE TRADITIONAL MODEL OF EVALUATION IN CONJUNCTION WITH STUDENT ACHIEVEMENT

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE SCHOOL OF EDUCATION

ATLANTA UNIVERSITY IN PARTIAL

FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF EDUCATION

BY

ANNE LUCILE GRIER CRAWFORD

ATLANTA UNIVERSITY
ATLANTA, GEORGIA
APRIL, 1986

ABSTRACT

TEACHERS' PERCEPTIONS OF THE GOAL-SETTING MODEL AND THE TRADITIONAL MODEL OF EVALUATION IN CONJUNCTION WITH STUDENT ACHIEVEMENT

Crawford, Anne Lucile Grier, Ed.D.
Atlanta University
Major Advisor: Trevor Turner, Ph.D.

The traditional model of teacher evaluation, which consists of several steps is used by 65% of the school districts in the United States. It has become the subject of much criticism and much debate because of its many weaknesses such as: (1) it promotes "Watchdog" attitudes; (2) it does not enhance instruction due to low teacher involvement; and (3) it causes the relationship between the principal and the teacher to deteriorate because evaluations are only used for administrative purposes. Thus, according to McGreal, the goal-setting model which consists of several steps is superior to the traditional model because, unlike the traditional model, it allows teachers to take an active role in the evaluation process, and it focuses on the individual needs of teachers.

This study attempted to determine whether the goal-setting model, based on the perceptions of teachers, was superior to the traditional model. In addition, it endeavored to determine whether the goal-setting model was superior to the traditional model in improving student achievement.

Based on the theoretical concepts and the empirical studies that were utilized in this study, the following research questions emerged:

1. Is there a significant difference between the teachers' perceptions of the goal-setting model and the traditional model and each of these in conjunction with student achievement?

2. Do teachers perceive the goal-setting model to be significantly different from the traditional model in:

   a. correcting weaknesses and enhancing strengths of teachers?

b. establishing a positive working relationship between the evaluatee and the evaluator?
c. emphasizing the professional growth and needs of teachers?

d. focusing on the self-evaluation of the teacher?

e. integrating individual performance objectives with the goals and objectives fo the school?

f. improving student achievement?

Two of the four middle schools in a large metropolitan area were selected to participate in this study. Sixteen of the teachers at each school and 15 of the students of each teacher were randomly selected to participate in this study. The questionnaires used in this study were the Teacher Evaluation Questionnaire and the California Achievement Tests.

This was an experimental investigation. The experiment was conducted for a period of six week, with the randomized group pretest, posttest design and the t-test was used to analyze the data.

The major findings for this study were:

1. There was a significant difference between the goal-setting model and the traditional model; however, there was no significant difference between the goal-setting model and the traditional model in improving student achievement.

2. There was a significant difference between the goal-setting model and the traditional model in the following areas: correcting weaknesses and enhancing strengths of teachers; establishing a positive working relationship between the evaluatee and the evaluator; emphasizing the professional growth and needs of the teacher; focusing on the self-evaluation of the teacher; integrating individual performance objectives with the goal and objectives of the school; and improving student achievement.

Some of the recommendations which enamated from this study were: (1) additional studies utilizing larger samples in more diverse areas and over a longer period of time should be conducted to replicate this study; (2) prinicipals should familiarize themselves with the current research or evaluation and utilize the research in their endeavors to improve teachers' performance; and (3) workshops on evluation should be held simultaneously for principals and teachers.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

Acknowledgements

The author wishes to thank the teachers and principals who participated in this study. Further appreciation is extended to the four-hundred and eighty middle school students who also participated in the study.

The writer is especially appreciative of the time, support, encouragement and advice given by Dr. Trevor Turner, the writers' advisor and chairperson of the dissertation committee.

Further appreciation is expressed to the other members of the dissertation committee-- Dr. Ganga Persaud and Dr. Olivia Boggs for their time, advice, expertise and encouragement to the writer of this dissertation.

The researcher would like to give special thanks to my alternative advisors Dr. Mollie Keaton and Dr. Loras Keaton for their invaluable help throughout this endeavor. I am also grateful to Dr. Fannie Hogan who gave me a great deal of support and encouragement during the preparation of this study.

Finally, the researcher must express her deep and profound appreciation to her friends for their understanding and support, and to her family for living with a part-time mother and sister. Their acceptance and realization of the importance of sharing the writer made it possible to dedicate this dissertation to Yasmin, Dianne and Ethel.

List of Tables

# CHAPTER I

## INTRODUCTION

There seems little need to offer extensive justification for the existence of teacher evaluation. Among educators, however, it is one of the few areas in which there is agreement. On the local level, for instance, educators agree about the espoused versus the "real" purpose of teacher evaluation. Moreover, on the national and state levels, educators agree that the overall purpose of evaluation is to safeguard and improve the quality of instruction received by students. Bolton[1] listed the following specific functions of teacher evaluation as the means for fulfilling this major purpose:

1. To improve teaching through the identification of ways that change teaching systems, teaching environments, and teaching behavior;

2. To supply information that will lead to the modification of assignments such as placements in other positions, and terminations;

3. To protect students from incompetence and teachers from unprofessional administrators;

4. To reward superior performance;

5. To validate the school system's selection process; and

---

[1] D. L. Bolton, Selection and Evaluation of Teachers (California: McCutchan, 1973), p. 5.

6. To provide a basis for teacher career planning and professional development.

If all of this agreement exists, why does teacher evaluation remain an extraordinarily controversial descriptive influence within local school settings? In most instances, the difficulties arise not with the concept or the general purpose, but from the way evaluation is carried out. Actual evaluation is most often directed by the requirements of the evaluation system. This creates problems, because in many cases the system is the problem.[2] This is not surprising because the major difficulties associated with teacher evaluation systems are well documented. They include such things as poor teacher-supervisor attitudes toward evaluation, the difficulties in formative and summative evaluation,[3] inadequate measuring devices, lack of reliable and

---

1 D. L. Bolton, Selection and Evaluation of Teachers, p. 5.

2 R. L. Wagoner and J. P. O'Hanlon, "Teacher Attitudes Toward Evaluation," Journal of Teacher Education 19 (1968):471-475.

3 Thomas L. McGreal, Successful Teacher Evaluation (Virginia: ASCD, 1981):vii.

consistent teaching criteria,[1] and the general lack of training of teachers and supervisors in evaluation.[2] Regardless of these difficulties, however, most school districts must have a functioning evaluation system. Whether the mandate is legislative, contractual, political or professional, the average school must be able to point to some systematic evaluation procedure to monitor the performance of its employees.

One systematic procedure used by 65% of the school districts in the United States is the traditional model. In most instances, no one usually claims credit for the development of this model, choosing instead to place the responsibility on some past anonymous committee.

According to McGreal,[3] the traditional model consists of the following steps:

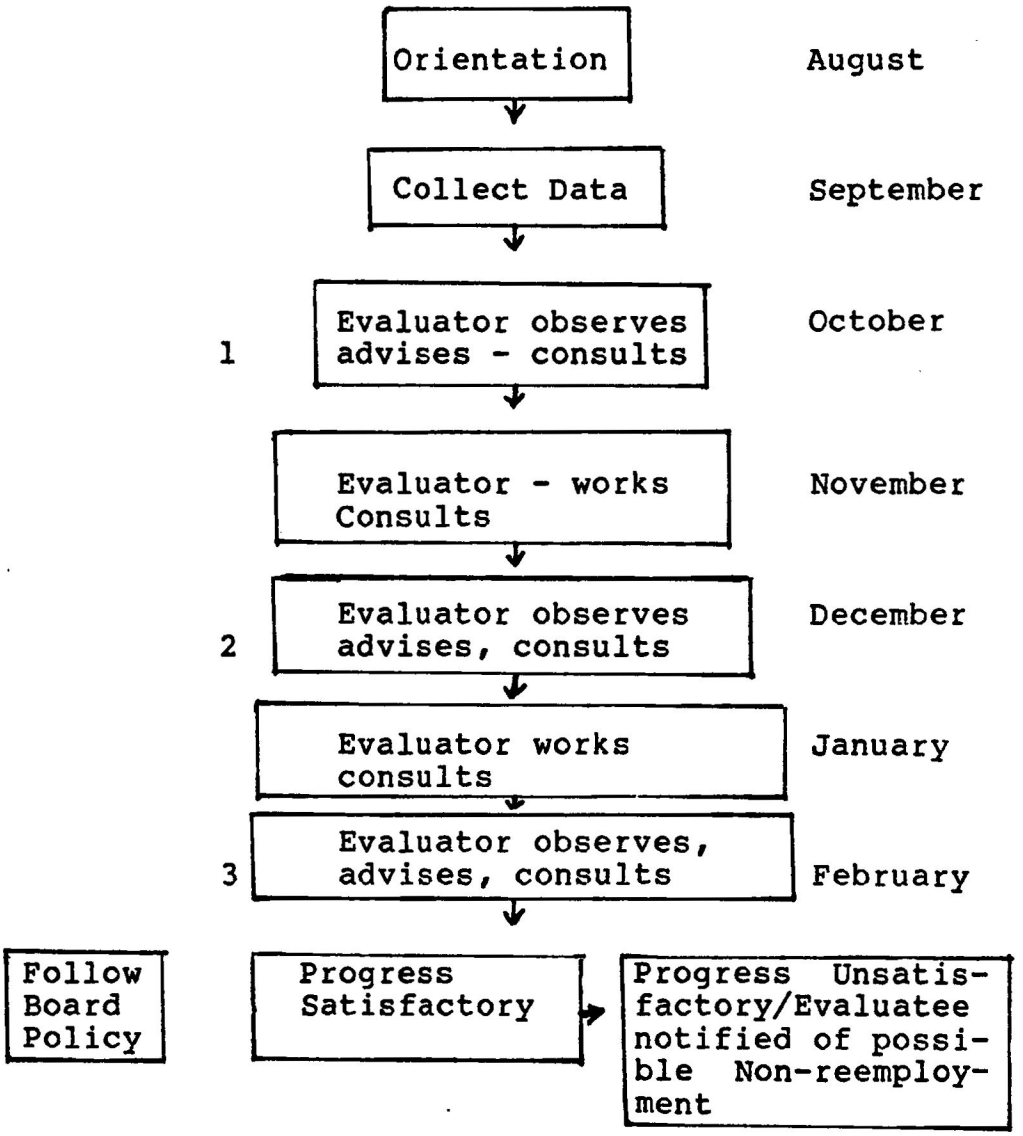
---

[1] W. J. Popham, "Teacher Evaluation or the Wrong Test for the Right Job." Paper presented at the Annual Meeting of The American Educational Research Association, Los Angeles, 1981.

[2] R. M. W. Trovers, "Criteria of Good Teaching," in The Handbook of Teacher Education, ed. J. Millman (California: Sage, 1981), p. 6.

[3] Thomas McGreal, Successful Teacher Evaluation, p. 12.

Figure 1. Traditional Model




The preceding steps in the traditional model may be explained as follows:

1. All non-tenured staff are evaluated by their principals at least three times during the school year. A professional evaluation form is submitted after after each evaluation. The final report is filed no later than the end of the first week in March.

2. All tenured teachers are evaluated by the principal or his/her designee at least once each school year. A professional evaluation report is submitted by April 15th,

3. A conference is held with the staff member following each evaluation. The completed evaluation report is reviewed with the staff member during the conference. Suggestions for improving areas marked fair or weak are made along with plans for any follow-up visits. Both parties then sign the report.

4. Teachers have the option to write comments about any part of the evaluation in the appropriate space. (See Appendix A).

## Statement of the Problem

This study attempted to determine whether the goal-setting model, based on the perceptions of teachers, was superior to the traditional model. In addition, it endeavored to determine whether student achievement improved when teachers were evaluated with the goal-setting model.

The traditional model of evaluation is used by many American schools; however, due to its many weaknesses, it has become the subject of much discussion and much criticism. Thus, the weaknesses of the traditional model are outlined below:

1. The traditional model of evaluation promotes "watchdog" attitudes. The very nature of the traditional system is summative. This

---

[1] Thomas McGreal, Successful Teacher Evaluation, p. 14.

summative emphasis has a tendency to promote the use of evaluative data gathered for administrative purposes. It has been shown that a teacher's attitude toward evaluation is a significant factor in the effectiveness of a system. Zelenak and Snider found that teachers who feel evaluation is used for administrative purposes tend to regard the process negatively.[1]

2. Traditional systems promote low teacher involvement and minimal contact time between supervisors and teachers. Standard procedures in the traditional system always have something being done to the teachers: however, teachers only change when they feel they are a part of the process that is designed to help them improve their instruction. Experience suggests that it is unlikely that teachers will be willing to change if they feel no ownership at all in the system.

3. There is a heavy emphasis on standardized criteria. It is particularly paradoxical that school districts praise themselves for their programs emphasizing individual differences in students while maintaining a teacher evaluation system that relies on standardized criteria. In effect they are saying that regardless of grade level, subject matter, ability levels of kids, experience, training, and physical setting, all teachers can be compared on the same set of criteria. This concept, perhaps more than any other, is seen as the major

---

[1] M. J. Zelenak and B. C. Snider, "Teachers Don't Resent Evaluation If It's for Improvement of Instruction," Phi Delta Kappan 55 (1974): 348-349.

roadblock to jointly developed cooperative activities between teachers and supervisors in the area of supervision and evaluation. There is a presumptuousness about the concept that anyone can identify a finite number of criteria that are so important that all teachers should be compared against them. As can be seen by looking at figure 2, Appendix, the criteria tend to address relatively general areas of competence, deal with ambiguous definitions, and address a number of characteristics or traits in which there is virtually no evidence to support their impact on children and what they learn in school. Perhaps this characteristic of the traditional model, more than anything else, offers the major reason for considering other alternatives.

4. Closely related to the preceding criticism is the fact that most criteria on the traditional model instruments tend to be administrative rather than teaching criteria. The rating scale offers a classic illustration of typical criteria. Often as much as 70 percent of the criteria contained on the traditional model evaluation instruments relate to administrative and personal concerns rather than to items that deal with the teacher's performance in the classroom. This means of selecting criteria for inclusion on an instrument only reinforces "watchdog" attitudes, promotes the notion of a system designed primarily for administrative purposes rather than for instructional improvement, and focuses time and energy on the part of the supervisors and teachers in relatively unproductive areas.

5. Finally, the traditional model forces supervisors to make judgments between people when there

> is no need to do so. The due process procedures defined by law regarding the dismissal of tenured teachers do not require a district to make comparative judgments between people.

While the traditional model does have several weaknesses as noted, it nevertheless has some advantages. They are:

1. The traditional model can be used in situations of high teacher supervisors ratios. The traditional model by design is quick. The normal requirements do not demand extensive contact between supervisor and teacher. Consequently, it is possible for a supervisor to complete evaluation requirements on a large number of teachers. By a general rule of thumb, whenever a supervisor is responsible for the annual evaluation of more than 20 certified staff, the traditional model clearly works to the supervisor's advantage.

2. The traditional model requires very little training on the part of the supervisors. The system requires a short start-up time and makes very few demands on supervisors. It obviously does not take any specific training to assist supervisors in providing high inference ratings to general standardized criteria.

---

[1] Bolton, Selection and Evaluation of Teachers, p. 96.

[2] Thomas McGreal, Successful Teacher Evaluation, p. 12.

> 3. Traditional law systems allow districts to visibly meet accountability demands while minimizing the often disruptive influence of evaluation. The nature of standardized criteria, high inference ratings, high supervisor involvement, and reliance on single and logical methods of collecting data present an evaluation model that is understandable to educators who serve on boards of education.[1]

Although there are some advantages in using the traditional model, it is interesting to note that none of the advantages relate to, nor are they designed to improve instruction. According to McGreal,[2] then, this model probably provides the most divided and the most negative images of teacher evaluation.

The displeasure with the traditional evaluation model provided a major impetus for designing other evaluation models, such as the goal-setting model. The goal setting model is an individualized approach to evaluation, such as the (MBO), the setting of goals in terms of expected results, working toward

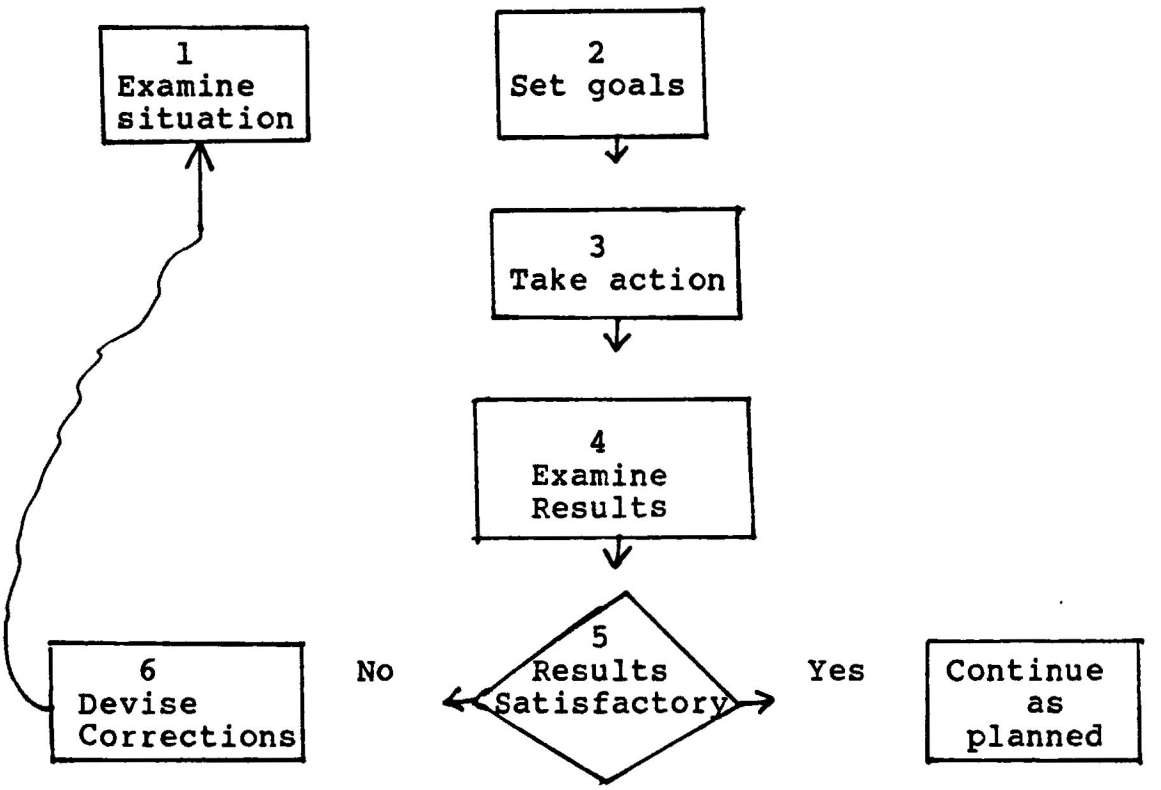
---

[1] Thomas McGreal, Successful Teacher Evaluation, p. 12.

[2] Ibid., p. 14.

these goals, and reviewing progress toward the goals. The flow-chart in figure 3 is a good illustration of steps that usually characterize goal-setting models.

Figure 3. Typical Procedures in Goal-Setting [1]




Iwanick provided basically the same outline of steps when he described the following procedures in the goal-setting approach:

1. Teacher conducts self-evaluation and identifies areas of improvement.

---

[1] Thomas McGreal, Successful Teacher Evaluation, p. 16.

2. Teacher develops draft of goal-setting contact.

3. Teacher and evaluator confer to discuss the teacher self-evaluation information, draft contracts, and the evaluator's perception of areas in which improvement is needed in an effort to reach agreement on the specifics of the contract for the current evaluation cycle.

4. Teacher and evaluator confer periodically to monitor progress toward goals stated in contract.

5. Teacher and evaluator confer near the end of the evaluation cycle to assess the extent to which goals have been accomplished as well as to discuss future directions for improvement which could be included in the goal contract during the next evaluation cycle.[1]

The pre-conference that begins the evaluation cycle clearly sets off goal-setting models from standard traditional procedures. In all of the various goal-setting models, the initial conference is viewed as the most valuable and the most important activity of the process. Hyman illustrated the importance of this step when he talked about the value of teachers and supervisors conferring together to develop goals. He stated that the initial conference does the following:

---

[1] E. F. Iwanick, "Conflict Plans," Handbook of Teacher Evaluation (Beverly Hills, California: Sage, 1981), p. 15.

1. Allows the teacher and the supervisor to explicitly focus their intentions in relation to the entire school context.

2. Requires the teacher and the supervisor to convene an initial meeting[1] to get to know each other better.

3. Requires the teacher and the supervisor to put their expectations in writing so as to have guidelines for future conferences, observations, and evaluations.

4. Offers an opportunity to talk about the improvement of teaching rather than[2] only maintenance of the status quo.

It is obvious that the goal setting model has many advantages. The advantages are as follows:

1. It promotes professional growth through correcting weaknesses and enhancing strengths.

2. It clarifies performance expectations and sets explicit criteria for evaluation.

3. It fosters a positive working relationship between teacher and evalutor.

4. It focuses on unique professional growth needs of each teacher.

5. It integrates individual performance objectives with the goals and objectives of the school organization.

---

1 R. T. Hyman, School Administrators Handbook of Teacher Supervision and Evaluation Methods (Englewood Cliffs, New Jersey: Prentice-Hall, 1975), p. 25.

2 Ibid., p. 20

6. It provides self-evaluation by the teacher.[1]

Even though the goal-setting model has several advantages, it also has certain weaknesses, such as:

1. It cannot be used to rank teachers.
2. It places too much emphasis on the attainment of measurable objectives.
3. It is not realistic in terms of the time and in-service resources available in most school settings.
4. It forces evaluators to make decisions about teachers' performance in areas in which they are not qualified.[2]

Although the goal-setting model has the aforementioned weaknesses, the literature seems to favor the goal-setting model over the traditional model because it is an individualized approach to evaluation with high teacher involvement; therefore, it will probably help to improve student achievement. Thus, this study will endeavor to determine whether teachers perceive the goal-setting model to be superior to the traditional model.

---

[1] Thomas McGreal, Successful Teacher Evaluation, p. 18.

[2] Ibid., p. 20.

## The Research Theory

This study attempted to determine whether the goal-setting model, based on the perceptions of teachers, was superior to the traditional model. In addition, it endeavored to determine whether student achievement is improved when teachers are evaluated with the goal-setting model.

According to McGreal,[1] the goal-setting model is superior to the traditional model because, unlike the traditional model, the goal-setting model allows teachers to take an active role in the evaluation process and it focuses on the individual needs of the teachers. This is important because, according to Gelzels and Guba,[2] when people are involved, they feel more committed to an organization and are therefore more enthusiatic. It would seem to appear,

---

[1] Thomas McGreal, Successful Teacher Evaluation, p. 25.

[2] Jacob W. Getzels and Egon G. Guba, "Social Behavior and the Administrative Process," The School Review 65 (Winter, 1957):423-441.

then, that student achievement will be higher when teachers are evaluated with the goal-setting model because, due to their personal involvement, teacher morale will probably be higher. In consensus with this statement, Kaura found that "student achievement increases with students who have teachers with higher morale and decreases with students who have teachers with low morale."[2]

---

[1] Bolton, Selection and Evaluation of Teachers, p. 20.

[2] Hussein S. Kaura, Student Achievement and Teacher Morale" (Ed.D. Dissertation, University of Michigan, 1963), p. 94.

## Research Questions

The questions which seemed most adaptive and likely to generate the informational requirements for the study were:

1. Is there a significant difference between the teachers' perceptions of the goal-setting model and the traditional model in conjunction with student achievement?
2. Do teachers perceive the goal-setting model to be significantly different from the traditional model in:
    a. Correcting weaknesses and enhancing strengths of teachers?
    b. Establishing a positive working relationship between the evaluatee and evaluator?
    c. Emphasizing the professional growth and needs of the teachers?
    d. Focusing on the self-evaluation of the teacher?
    e. Integrating individual performance objectives with the goals and objectives of the school?
    f. Improving student achievements?

## Definitions

The following definitions were operationally defined in this study:

1. Goal-setting Model - An evaluation procedure which emphasizes an individual approach to teacher evaluation.

2. Traditional Model - An evaluation system which relies on simplified definition and evaluation or procedures and processes that have remained unchanged for years.

3. Perception - Cognitive judgement of the goal-setting and traditional models of teachers.

4. Evaluation - Assessment of teachers' performance.

5. Control Group - Used interchangeably with traditional group and group 1.

6. Experimental Group - Used interchangeably with goul-setting group and group 2.

7. Teacher - Person who has been assigned to a middle school to teach academic/non-academic subjects and perform other duties stipulated by the principal.

8. Superior Evaluation Model - Model receiving highest rating from teachers.

9. Middle School - Schools with grades 6-8.

10. Professional Growth - Knowledge gained by taking responsibility for accepting and supporting the importance of teachers' development in areas such as classroom management.

11. Performance Expectations - Expectations regarding performance of teachers held by supervisors, including a teacher who is knowledgeable about child growth/development, plans learning activities in accordance with individual differences, knowledge of subject matter, teaching methods techniques and instructional skills.

12. Working Relationship - The pattern of behavior established between teachers, students, peers, administrators, supervisors, and parents.

13. Self-Evaluation - The teacher's own assessment of his/her performance in classroom management, such as establishing good relationships.

14. Student Achievement - Student's score on the California Achievement Tests.

## Significance of Study

The significance of this study lies in its potential for providing a basis on which public school districts may develop viable and defensible procedures for the evaluation of teachers. Beyond this, the research should provide a source, through

the review of the literature, for anyone seeking information on the present state of thought and practice in the area of teacher evaluation. Finally, it is anticipated that the drawing together of prominent aspects of research and writing in this area will help to clarify its present developmental stage and to emphasize some of the elements most in need of refinement and further study.

## Limitations

The following limitations controlled this study:

1. This study was confined to one geographic area in Georgia; therefore, the only generalization possible is to the population from which the sample was drawn.
2. The data for this study were based on the perceptions of teachers.

## Assumptions

The following assumptions were made at the beginning of the study:

1. Teacher behavior has an effect upon student behavior.
2. There exist in schools expectations, stated or implied, for student behavior and hence, for teacher behavior.

3. In light of these expectations, there are some teacher behaviors that are more desirable than others.

4. Teacher behaviors can be modified to make them more desirable in light of the expectations of the school.

5. Supervisory behavior that is guided by a well defined procedure is an effective means of modifying teacher behavior.

## Organization of Study

Chapter I was an introduction to the study which included a problem statement and research questions, significance of the study, a research theory, assumptions, limitations and definitions. Chapter II is a review of the research methodology. The findings are presented in Chapter 4, and a discussion of the findings, concluding statements and recommendations are found in Chapter V.

# CHAPTER II

## REVIEW OF RELATED LITERATURE

The review of the literature is divided into two sections: a review of theoretical concepts, which are discussed under the heading, Processes of Teacher Evaluation, and includes classroom observation, self-appraisal, management by objectives (MBO) and student evaluation of teachers. The empirical studies are presented under the following captions:

Development/Analysis of Instruments for the Evaluation of Teachers

Models for the Evaluation of Teachers

Evaluation of Teacher Competency

Self-Evaluation Models

Factors which Influence the Principal's Evaluation of Teachers

Overall, the literature suggested and supported the idea of a revised method of teacher evaluation.

### Review of Theoretical Concepts

The theoretical framework focused on various teacher evaluation methods. These methods are discussed in the following section under "Processes of Teacher Evaluation."

### Processes of Teacher Evaluation

The process of teacher evaluation can be viewed as belonging to four categories: 1) classroom

observation, 2) self-appraisal, 3) management by objectives (MBO), and 4) student evaluation of teachers.

The first process, classroom observations, is implicative of an attendant set of supervisory behaviors and interactions between teacher and evaluator that commonly occur in the form of conferences before and after observation. The self-appraisal process takes several forms, but all share the element of the teacher in some way viewing his/her behaviors in relation to some predetermined set of criteria. The MBO approach represents the selective application of concepts embodied in basic goal setting and attainment measurement. All of the examples of student evaluation encountered in the literature utilized a structured questionnaire which elicited a value-laden response from the student.

Dividing the processes of teacher evaluation into four categories is not meant to suggest that given processes do not share elements of more than one category, since most do. The identification of the four categories does, however, facilitate discussion of the numerous practices and variations encountered. Thus, the discussion that follows is intended to characterize each of the four categories by examples found in the literature and by reference

to salient elements of the examples that could have potential significance in the identification of a set of criteria which may be employed for the development of a new procedure for teacher evaluation.

## Classroom Observation

The observation and rating of a teacher's classroom behavior reflects the operational modes of most elementary and secondary schools in which the classroom is viewed as the primary site of learning. As such, it is common to the extent of being inherent in teacher evaluation procedures.

The process of classroom observation operates on the premise that observable teacher behaviors are significant in determining instructional effectiveness and that for the most part, the evaluator is capable of making determinations of varying degrees of proficiency based upon his/her observations. The Commission of Public School Personnel Policies in Ohio, 1972),[1] The National Education Association

---

[1] "Planning the Evaluation of Educators," Washington, D.C. NEA Division of Instruction and Professional Development, 1972.

(1972),[1] Popham (1971),[2] Hickox (1976)[3] and Iwanicki (1981)[4] do not argue with this premise, but all call attention to the fact that the evaluator judgements lack reliability and validity. The NEA and Drumheller (1974)[5] both maintain that the personal values held by the observers are the prime determinants of the ratings given. Popham refers not only to the absence of a common agreement of what is good in teacher classroom behavior, but also to the impossibility of developing general indices of teacher skills because of the variability of the observation situations.

---

1 "Where They Not Only Evaluate Teachers, They Actually Help Good Ones and Get Rid of Bad Ones," American School Board Journal. 163:10, (1976).

2 W. J. Popham, "Teacher Evaluation". Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, 1981.

3 Edward Hickox. "The Shape of Teacher Evaluation: A Survey of Practices in the Capital District of New York" (Albany: State University of New York, 1976).

4 E. Iwanicki, Handbook of Teacher Evaluation (California: Beverly Hills, California Publication, 1981).

5 Sidney J. Drumheller, "Evaluating Teachers Through a Jaundiced Eye," Educational Technology. 14:7 (1974), pp. 17-22.

Tuckman (1977),[1] in his study of inter-rater reliability, found that the ratings across all principal groups, elementary, intermediate and senior high were not reliable; however, intra-group reliability did exist. Senior high principals were in agreement on one set of desirable behavioral characteristics for teachers; however, the intermediate principals were in agreement on a second set of characteristics and the elementary on a third set.

The sophistication and complexity of observation processes reviewed vary greatly from systematic use of Flanders Interaction Analysis Scale for charting behavior (Hull and Hanson, 1972),[2] to a simple checklist of general characteristics such as appearance, knowledge of subject matter and classroom environment (Stemrock, 1972).[3] Although the use of

---

1 Bruce W. Tuchman, "Feedback and Change Process," Phi Delta Kappan, 57:5 (1976):341-344.

2 Ray Hull and John Hanson, "Classroom Supervision and Informal Analysis of Behavior," A Manual for Supervision (Oveson: Oregon School Study Council, 1972).

3 Suzanne Stemrock, Evaluating Teachers Performance (Washington, D. C. Educational Research Service, 1972).

the checklist-type of instrument predominates, McKenna (1971),[1] Ryans (1971),[2] Poliakoff, (1972)[3] and Goldstein, (1982)[4] in their reviews of evaluation practice, all rate a definite trend away from checklists to procedures that are more specific and more objective.

Whatever criticism is made of the existing practices in classroom observation, no author encountered in the literature advocated its abolition as a means of evaluations. Rather the extent of the discussion of the process implies an acceptance of its necessary existence.

## Self-Appraisal

Self-appraisal is being increasingly incorporated as a part of evaluation procedures (Poliakoff, 1973). Of the self-appraisal processes

---

1 Bernard H. McKenna. Teacher Evaluation: An Annotated Bibliography (New York: McGraw, 1971).

2 David G. Ryans, "Teacher Evaluation Research Part I: Consideration of Critical Issues, Feasibility of Collaborative Research and Overall Design, Final Report" (Honolulu: ERDC Center, 1971).

3 Lorraine L. Poliakoff, Evaluating School Personnel Today (Washington, D. C.: ERIC Clearing House on Teacher Education, 1973).

4 R. Goldhammer, Clinical Supervision (New York: Holt, Rinehart, and Winston, 1969).

described in the literature, the Batelle Self-Appraisal Instrument (Ohio Education Association, 1970)[1] is the most notable. Unlike most other self-appraisal processes reported, it focuses on the teacher's behavior outside the classroom as well as within the classroom. The instrument causes the teacher to examine himself/herself in the roles of 1) instructional leader, 2) developer of self concepts, 3) promoter of healthful emotional growth, and 4) communicator with parents and colleagues.

The use of audio and video tapes in conjunction with the interaction analysis scales of Roberson and Flanders are representative elements of self-appraisal processes (Oldham, 1972).[2] The different instruments have in common the comparison of teacher perception and behavior to some pre-determined standard or quality.

The basis for the contention that self-appraisal is an effective means of improving instruction lies in the assumption that a teacher's behavior can be

---

[1] Teacher Evaluation: Interface on Learning (Ohio: OEA, 1970).

[2] Neil Oldham, "Evaluating Teachers for Professional Growth: Current Trends in School Policies and Programs," An Education U.S.A. Special Report (Virginia: NSSPA, 1974).

altered by the individual without outside direction -- that true behavioral change occurs only when the individual acknowledges and internalizes discrepancies in his/her behavior through his/her own volition.

## Management by Objectives

A management by objectives process also referred to as planning, review and job target identification, is the most recent of the four types of processes to develop. MBO, which had its origin in industry, is now being presented to schools as the answer to some of the prevailing accountability questions. While its use is still somewhat limited, it is becoming more widely accepted and applied (Poliakoff 1973).[1]

The Redfern model (Ohio Education Association, 1970)[2] is the most comprehensive procedure described in the literature reviewed. While it includes components from several of the four categories described, it is primarily an MBO process. The Redfern model treats the total teaching job as consisting of five components: 1) classroom instruction, 2) consultation with individual pupils, 3) effective communication, 4) professional participation, and 5) self-criticism and analysis. The plan seeks to evaluate all four of these

components by a total process marked most by frequent and continual contact between teacher and supervisor. The total process required 1) definition of the individual teacher's job, 2) identification of major areas of responsibility, 3) designation of specific job objectives, 4) explanation of the relationship between supervisory activities and evaluation, 5) evaluation of the teacher by the supervisor, 6) an evaluation conference between teacher and supervisor, 7) provision for follow-up communication between teacher and supervisor, and 8) teacher self-evaluation.

The Redfern plan also requires, as do some other such MBO type process, (Teacher Evaluation Guide Atlanta Public Schools, 1982) that the teacher and supervisor reach mutual agreement on the selection of job objectives as well as the strategies to achieve them. Additionally, in many cases the teacher is the initiator of objectives with the supervisor behaving as reactor. It is these characteristics of MBO type processes that draw criticism such as theat of Geiser (1973)[1] and the Commission of Public School Personnel

---

[1] Thomas L. McGreal, Successful Teacher Evaluation, p. 18.

Policies in Ohio (1972).[1] Geisert contends that these characteristics place the supervisor in a defensive position where it is up to him/her to establish that the teacher initiated objectives and strategies are inappropriate. It also places him/her in a negotiating position which would appear to leave little room for supervisory direction for improving teacher performance.

## Student Evaluation of Teachers

Poliakoff (1973)[2] found little evidence of student evaluation of teachers at the high school level. Instances of its occurrence noted by her were primarily adaptations of processes utilized by colleges. There is no evidence of use below the middle school level.

One instrument being used is the Teacher Image Questionnaire prepared by the Educator Feedback Center, Western, Michigan University. The questionnaire solicits student response, on a five point scale, to 16 teacher characteristics as well as

---

[1] Teacher Evaluation: Interface on Learning (Ohio: Ohio School Association, 1970).

[2] Lorraine L. Poliakoff, Evaluating School Personnel Today.

providing two queries on teacher strengths and weaknesses.

The Illinois Course Evaluation Questionnaire (Aleomoni, 1973)[1] is a college level instrument, but may have applicability to the high school level if somewhat modified. It solicits students' responses on a Likert scale to 19 different categories of teacher and course characteristics. The most notable aspect of this questionnaire is the extensive bank of data that has been developed which supports its reliability and validity. The CEQ is recommended for use by the individual teacher, and its results should be shared only in the context of a larger multi-faceted system of evaluation.

The National Institute of Education has developed a Teacher Evaluation Questionnaire (Melnyk, 1973)[2] for use in secondary schools. The students rate the instructor on personal characteristics and professional characteristics, but as distinctly

---

[1] Lawrence M. Aleomoni, Evaluation by Students to Identify Instructional Problems (Urbana, Illinois: Office of Instructional Resources, Illinois University, 1973).

[2] M. Melnyk. The Teacher Evaluation Questionnaire and the Quantification Results (Washington, D. C.: U.S. Department of HEW, National Institute of Education, 1973).

separate from one another. The responses are measured against a set of criteria which supposedly represent the behavioral profile of a "good teacher."

In researching student evaluation of teachers, Jackson and Kenny (1972) and McGreal (1980)[1] discovered some things that may emphasize a need for caution in the use of student evaluation of teachers. Jackson and Kenny found that students of middle class socio-economic background rate teachers significantly higher than do students from lower class backgrounds. Students saw teachers who shared with them the same socio-economic background as more effective in communication than teachers of a different background.[2]

McGreal (1980) found that in testing a student's perception of the teacher in a value-laden (good-bad) way, the student, regardless of his level in school (middle school, high school or college freshman) will identify as good those things which he/she like or feels good about. For example, middle schoolers rated the teachers highest on the amount of recess time allowed and high school students valued a sense of humor in the teacher above all other

---

[1] Thomas L. McGreal, Successful Teacher Evaluation, pp. 28-30.

[2] Ibid., p. 151.

characteristics.

Despite the tenuous value of student evaluation of teachers implied by the findings of Jackson, Kenny and McGreal,[1] and the NEA Division of Instruction and Development (1980)[2] maintains that its examination of research on the topic reveals a low but significant correlation between student rating of an instructor and student achievement.

Poliakoff (1973)[3] also points out some advantages to the process of student evaluation of teachers, particularly if the students focus only on observable behaviors. She points out that students have a longer time to observe a teacher in many more different situations than would a supervisor. She also calls attention to the fact that through student rating, a large number of individual biases can be averaged -- an impossibility when a single supervisor does the evaluation. Contradictory to the NEA findings, Poliakoff maintains that her search found no evidence

---

1 Planning the Evaluation of Educators: A Simulation Handbook for Conducting a Simulation (Washington, D. C.: NEA Division of Instructional and Professional Development, 1972).

2 Lorraine L. Poliakoff, Evaluating School Personnel Today.

3 Ibid., p. 18.

of a significant relationship between student ratings and teacher performance.

Wilcox (1976)[1] states that senior high students' appraisal of teachers did not correlate with the principal's appraisal nor did it correlate with the appraisal of some teachers by recent graduates.

## Review of Empirical Studies

The researcher also reviewed several relevant empirical studies. These studies are presented in the succeeding section.

## Development/Analysis of Instruments For the Evaluation of Teachers

Leeth[2] sought to develop and validate an industrial arts teacher rating instrument.

A comprehensive list of industrial arts teacher rating items was compiled through the review of literature and solicitation of rating instruments from Texas supervisors of industrial arts. A teacher pool

---

[1] Ray T. Wilcox, "A Comparison of Secondary School Teachers Judged Effective by Principals, Current Students, and Graduates." Paper presented at the annual meeting of the California Educational Research Association (55th Burlingame, Californai: November 18-20, 1976).

[2] Benton M. Leeth, "The Development and Validation of an Industrial Arts Teaching Rating Instrument" (Ph.D. dissertation, Texas A&M University, 1977).

was identified by soliciting cooperation from sixteen supervisors who provided the names and addresses of 600 teachers.

A teacher rating instrument consisting of 211 items was developed and submitted to a pilot group of 50 industrial arts teachers. Using the mean response scores from the pilot group, the instrument was revised and reduced to 174 items for submission to the primary teacher input sample. Returns from this group were used to select the upper quarter of the rating items for retention in the validation instrument.

The correlation coefficient, multiple regression and analysis of variance were utilized to analyze the data. Among the major findings were: (1) teacher selected rating items are acceptable to supervisors; (2) teacher selected rating items are good predictors of teaching success as compared to supervisor's ratings; and (3) sophisticated analysis of rating instruments is a useful technique for identifying the most valid and most predictive items and for shortening the length of the instrument.

Adams[1] also sought to develop and field test an

---

[1] Gwendolyn J. Adams, "Development and Field Testing of an Evaluation Instrument for Assessing Competencies of Teachers of Learning Disabled Students" (Ph.D. dissertation, Auburn University, 1977).

instrument; however, the observation instrument was to be used in recording observable behaviors relative to five global competencies judged important for teachers of learning disabled students in resource classes. The five global competencies were the learner's objectives, motivation, learning theory, meeting the needs of the individual students and classroom management. Each competency was followed by a list of statements designed to describe observable behavior in the classroom. Observers recorded on the instrument the observed behaviors.

The sample was composed of all 16 learning disabilities teachers assigned to resource classrooms in the Montgomery, Alabama Public Schools. Seven individuals, recognized as authorities in the field of learning disabilities, were asked to rate the descriptive statements having high, moderate, or low ability to indicate the presence of the competency. Five completed ratings were returned and incorporated into the study.

The results of the field test of the instrument indicated that observers were in agreement in recording behaviors observed; that instruments reflected mastery of stated competencies by most teachers in the sample; and that school principals recorded observation of more behaviors than did other

observers.

Inglis[1] conducted a study which was threefold: (1) to develop and validate an instrument to measure student's reports of perceived teacher invitations and perceived teacher behaviors traditionally considered effective by authorities; (2) to subject the instrument to statistical procedures for validation; and (3) to correlate identified invitational (I-type) factors and effective (E-type) factors with academic achievement, determined by student reports of grades earned, how hard they worked and how much they learned.

The working model was illustrated by reviewing the current closure of two schools in Tulsa, Oklahoma where the model was applied. The instrument/scale identified five discriminate factors which the investigator labeled "Caring," "Respect," "Course Organization," "Interpersonal Contact," and "Learning Environment." The first two factros, caring and respect, identified I-type teacher behaviors. The final three factors, course organization,

---

[1] Sandra Inglis, "The Development and Validation of an Instrument to Assess Teacher Invitations and Teacher Effectiveness as Reported by Students in a Technical and General Post Secondary Educational Setting" (Ph.D. dissertation, University of Florida, 1976).

interpersonal contact, and learning environment, identified E-type teacher behaviors.

The results were: (1) there is a significant positive relationship between I-type and E-type teacher behaviors as perceived and reported by students; (2) there is a significant positive relationship between teacher behaviors (I-type and E-type) and student reports of grades earned in the course; (3) there is a significant positive relationship between teacher behaviors (I-type and E-type) and student reports of how hard they worked in the course; (4) there is a significant positive relationship between teacher behaviors (I-type and E-type) and academic achievement on the part of the students.

The purpose of William's[1] study was to analyze teacher evaluation instruments used in the Tennessee Public School Systems relative to type of instrument and content in relation to the characteristics of six school systems. Differences were sought between school systems in relation to participation in teacher contract negotiations, per pupil expenditure, average

---

[1] Jane L. Williams, "Analysis of Teacher Evaluation Instruments in Use in Tennessee Public School Systems" (Ph.D. dissertation, East Tennessee State University, 1983).

annual teacher salary, method of selection of superintendents, type of school system, and size of school system in their choice of type of instrument and in the content of their instruments.

A content analysis coding sheet was devised to use in the study categorizing the content of teacher evaluation instruments as: (1) personal qualities references, (2) professional qualities references, (3) instructional skills references, and (4) classroom management references. Instruments were typed as (1) rating scales, (2) checklists, (3) performance objectives, (4) anecdotal records, and (5) combination rating scales and performance objectives.

Teacher evaluation instruments from 129 Tennessee Public School Systems were analyzed and coded for type of content. The chi-square test was used to determine differences between school systems in their choice of type instruments.

The findings of the study were: (1) the overwhelming majority of Tennessee school systems used some form of rating scale in the evaluation of teachers; (2) a significant difference at the .05 level was found between school systems with average annual teacher salary below a nd above the median in their choice of type of instrument; (3) a significant difference at the .005 level was found between county

and city/special school systems in their choice of type of instrument; (4) a significant difference at the .04 level was found between school systems with average annual teacher salary below and above the median in percentage of instructional skills references on their evaluation instruments; and (5) a significant difference at the .008 level was found between county and city/special school systems in the percentage of instructional skills references on their evaluation instruments.

Shirer's[1] study addressed the question: What purposes for school evaluation do educators perceive as desirable and what purposes for school evaluation do educators perceive as forced upon them?

One person from each of four role groups -- teachers, principals, superintendents, and school board presidents associated with school accredited by the North Central Association in Wisconsin -- constituted the study population. A total of 843 persons were mailed questionnaires. The response was 66.7 percent. This non-random sample was

---

[1] William Roe Shirer, "The Purposes of School Evaluation as Perceived by Teachers, Principals, Superintendents and School Board Presidents of Schools Holding Membership in the Wisconsin North Central Association" (Ph.D. dissertation, The University of Wisconsin.

representative of the population from which it was drawn.

To determine the extent to which respondents perceived certain evaluation objectives as desirable and certain evaluation objectives as forced upon them, an instrument entitled, The Evaluation Goals Inventory (EGI), was designed. The instrument included theoretical categories of evaluation, accountability, decision-making and instructional improvement.

Factor analysis was employed to reduce the original seventy dependent variables to twelve empirical factors to be used in subsequent analysis. One-way analysis of variance was employed to determine the extent to which differences among respondents were associated with their role, and the size, type and setting of their school. Post hoc tests were then performed to further specify differences among groups.

Twelve factors representing categories of evaluation purpose were identified and named as a result of the factor analysis. Nine factors, different from the theoretical factors, represented evaluation purposes educators might want to pursue while three factors represented evaluation purposes educators might feel forced to pursue. Role-related differences among respondents were found by nine of

the twelve factors. Of the thirty-six predicted differences among groups of respondents, twenty-three were found to exist in fact. No significant difference among respondents were found with respect to any other background or descriptor variable employed in the analyses.

In his study, Schmitt[1] endeavored to develop a methodology for comparing different types of teacher evaluation systems, the costs involved with each, and the relative level of personnel satisfaction with each system. Rating scales, job targets, and multiple evaluator procedures were identified in different sized Iowa high schools. Total developmental costs and annual operational costs of evaluations were gathered for each school and teacher, administrator, and board member completed attitudinal questionnaires.

Among the findings were: (1) teachers were lease inclined to believe that improvement of instruction was the major purpose of evaluation; (2) teachers indicated the least amount of involvement in developing their evaluation procedure; (3) teachers were most critical of their evaluation procedures; (5)

---

[1] Edward L. Schmitt, "A Cost Effectiveness Analysis of Selected Teacher Performance Evaluation Systems" (Ph.D. dissertation, Iowa State University, 1976).

teachers held the lowest view of the board's priority of evaluation; (6) teachers held the lowest opinion of the worth of evaluation costs; (7) teachers held the lowest estimation of evaluation as the key element of systematic personnel procedures; (8) teachers least agreed that evaluation improves instruction; (9) large schools spent significantly more than medium-sized schools for both total developmental cost and annual per teacher cost of evaluation; and (10) no significant cost differences were found among the three types of evaluation procedures.

Heuss[1] sought to identify the administrative procedures, criteria, and instruments used to evaluate teacher performance and effectiveness in Texas public schools; to compare evaluation procedures instruments, and practices with theoretical recommendations; to propose a model for future use in instructional evaluation; and to solicit administrative recommendations for the improvement of teacher evaluation in Texas public schools.

A questionnaire was used to survey 1102 Texas public school superintendents to determine

---

[1] Ronald Neal Heuss, "Teacher Evaluation Purposes, Procedures and Instruments in Texas Public Schools" (Ph.D. dissertation, Baylor University, 1984).

administrative procedures, criteria, and instruments used to evaluate teachers. Seven hundred and nineteen superintendents, or 65.3 percent, participated in the study.

The findings were as follow: (1) most superintendents do not require training for teacher evaluators; (2) most superintendents agree the purpose for teacher evaluation should be to improve instruction; (3) most superintendents require annual summative evaluations without formative procedures recommended by teacher evaluation library authorities; (4) most superintendents desire less subjective criteria and fewer instruments; and (5) most superintendents desire additional techniques and most instruments to improve their teacher evaluation practices.

The primary objective of Janney's[1] study was to establish a set of guidelines for teacher evaluation systems. Criteria in the area of board policies, goals, the process of developing and implementing a teacher evaluation system and procedures used in evaluating teachers were derived from the evaluation

---

[1] Patra Moss Janney, "Teacher Evaluation Guidelines Based on Expert Judgement, National Laws, and State Laws of West Virginia" (Ph.D. dissertation, Virginia State University, 1984).

literature. Professional specialists throughout the United States indicated their agreement or disagreement with these criteria. Federal and state laws and regulations were reviewed to establish legal criteria for teacher evaluation programs. The largest number of inadequacies were found in the development and implemtation areas; the fewest inadequacies were found in the procedures area.

## Models for the Evaluation of Teachers

In this study, Spears[1] purported to develop guidelines for a program of teacher evaluation.

Data were collected by extensively surveying the literature in the area of teacher evaluation in order to answer the following questions: (1) What are the purposes of teacher evaluation? (2) What are the characteristics of a successful evaluation program? (3) What criteria should be included in a teacher evaluation program? (4) Who should be involved in developing the teacher evaluation program and who should administer and apply the teacher evaluation program?

The main purposes of evaluating teachers seem to

---

[1] Voy Spears, "Guidelines for Developing a Program of Teacher Evaluation" (Ph.D. dissertation, The University of Arkansas, 1974).

be for the improvement of instruction, teacher accountability, due process of law, negotiations, teacher tenure laws, the taxpayer's dividends and to give teachers an accurate assessment of their own performance in the classroom.

Based on his findings, Spears concluded that representative members from all areas of the educational community should be involved in the development, implementation, fuction, evaluation and updating of teacher evaluation programs, and that evaluation programs with a competency-based program of performance objectives are more successful. He stated further that the succeeding evaluative criteria should be included in an evaluation program: coopertively developed performance assessment of what students learn, and a method for assessing some of the teacher's routine duties and the teacher's ability to control and protect the health and safety of the stuents while they are in school. All factors of appraisal criteria should be cooperatively developed involving the evaluator, the evaluatee and the representative members of all those who have a vested interest in the evaluative process. Principals, representative members from all levels of administration, peer groups, supervisors, parents, students, and teachers should be involved in

developing, administering and applying the teacher evaluation program.

Brungardt[1] also endeavored to present a more feasible and effective way of approaching the problem of evaluation of teachers in a public school setting by suggesting selected means that are reflective of recent thought and practice regarding the topic.

The model for evaluation presented was designed to allow for variations in expectations from district to district, and for the accommodation of compromise among the affected parties. Based upon a currently employed procedure for evaluation and modified to acknowledge recent research and practice the model includes a format for a development process, the identification of necessary components of an evaluation procedure, a description of the mechanics of implementation, recommendations for the in-service education of evaluators and evaluatees, a process for continual reassessment of an evaluation procedure, and sample forms for documenting the evaluator/evaluatee interactions.

Using recent research and continuing practice as

---

[1] Karl Brungardt, "A Procedural Model for the Improvement of Instruction through the Evaluation/Supervision of Teachers" (Ph.D. dissertation, The University of Minnesota, 1979).

defense, the author offered five assumptions on which the systematic evaluation of teachers may be based:

1. Teacher behavior has an effect upon student behavior.

2. There do exist, in schools, expectations for student behavior and hence, teacher behavior.

3. In light of these expectations, there are some teacher behaviors that are more desirable than others.

4. Teacher behaviors can be modified to be more desirable, i.e., consonant with the expectations of the school.

5. Supervisory behavior that is guided by a well-defined evaluation procedure is effective means of modifying teacher behavior.

Whereas the preceding researcher endeavored to present an effective model for teacher evaluation, Willard[1] assessed a teacher evaluation model based on McGregor's Theory "Y" assumption, to determine if such an approach to teacher evaluation could produce statistically significant positive changes in the attitudes of teachers and supervisors toward evaluation.

The teachers and supervisors from the University

---

[1] David T. Willard, "An Assessment of the Effects of a Staff Evaluation Model Developed from Douglas McGregor's Theory "Y" Upon the Attitudes of Teachers and Supervisors Toward Evaluation" (Ph.D. dissertation, St. Louis University, 1979).

of St. Louis developed a special questionnaire to measure the attitudes of teachers and supervisors about teacher evaluation. All supervisors and a random sample of teachers in an immediate school district in a suburban county (Midwest) mailed a questionnaire prior to the implementation of the positive approach to evaluation developed by the committee based on McGregor's Theory "Y" assumptions. Questionnaires were mailed to all supervisors responsible for teacher evaluation and another random sample of teachers after the new approach had been in use for one year. The BMPDV program of analysis of variance and co-variance was utilized to analyze the data.

Among the major findings were: supervisors expressed significantly greater agreement with seven statements concerning evaluation which were positively related to McGregor's Theory "Y" assumptions and only one item which was related to Theory "X" assumption. Teachers expressed significantly greater agreement with items concerned with evaluation as a means of controlling the teacher's work and eevaluation as a proces completed to comply with tenure laws. Both teachers and supervisors indicated stronger agreement on post-test items: evaluation is to encourage professional growth and evaluation helps a teacher to

determine in-service needs. Supervisors indicated greater agreement with the items. Evaluation is done for a teacher and not to a teacher and evaluation helps determine in-service needs.

In his evaluative study, Heebink[1] sought to provide an urban Wisconsin school district's dicision-makers with information and recommendations relative to the value of a developmental teacher evaluation model, especially as it affected the effectiveness of experienced elementary school teachers.

To organize the considerable data gathered for this comprehensive investigation, the writer used the substance of a 3x4 matrix proposed by Robert Stake in 1967. The model's three rows were antecedents, transactions, and outcomes. Intendents, observations, standards and judgments were the column titles. Two comparable Wisconsin school systems help establish relative standards.

The observation accuracy of principals using the developmental program was highly rated by experienced elementary teachers. Target district principals had also acquired more ideas for improving instruction

---

[1] William Heebink, "A Comprehensive Investigation of a Developmental Teacher Evaluation Program's Effectiveness with Experienced Elementary Teachers in an Urban Wisconsin School District" (Ph.D. dissertation, Florida State University, 1979).

from their evaluation program than their peers had from the comparison programs. In terms of time demands, the developmental program appeared practical for principals. More target district personnel viewed their evaluation program asaimed at instructional improvement than was the case in either "control" district. Experienced teachers preferred the target program to a product program alternative.

Swanson[1] attempted to identify which mode of teacher evaluation was most efficient, based on farily objective performance criterion, and to establish a basis for viewing teaching style as leadership style. In existing research, superior ratings were the most used evaluation measure, student ratings were a rapidly growing mode of evaluation, self-ratings were considered biased, and peer ratings were used very little. Hence, who should do the evaluating was an unsolved problem. All four evaluation modes were employed in the present study for comparison.

A number of current teacher evaluation questionnaires were examined to ascertain what educators felt constituted good teaching. Every teacher evaluation questionnaire was paralleled to one or more supervisory behavior description among which leader behavior varies: consideration and structure. Every teacher evaluation questionnaire item was

parrallel to one or more SBD items. An educational setting was selected from the U.S. Air Force pilot training system. Five different performance aspects were weighted and combined to form a multiple-input criterion which included an instructor's own performance as well as the performance of his students. The two SBD dimensions and a forced-ranking were used as three predictor variables. They were combined through the use of a multiple linear regression model.

The findings indicated that peers had significantly less prediction error than all other groups. Superiors were next, followed by self and students.

Since clinical supervision has been established as an effective model for evaluation,[2] Scime investigated whether there was a significant difference between the perceptions of teachers and building administrators regarding the extent that the

---

[1] Ronald George Swanson, "Teacher Evaluation as a Function of Leadership Style: A Multiple Correlational Approach" (Ph.D. dissertation, North Texas State University, 1974).

[2] Joseph Scime, "The Conflict Between Evaluation and Supervision: A Study of a Clinical Supervision Program Designed to Link Instruction, Supervision, and Evaluation of Staff Development" (Ph.D. dissertation, George Peabody Teachers College, 1984).

objectives of a program of clinical supervision were being met. Also, this study attempted to determine whether any differences that existed could be attributed to specific variables such as professional role, tenure status, frequency of observation, and academic department.

To determine how the 595 teachers and 55 administrators felt about the clinical supervision program, 14 research questions were identified and translated into statements for a survey questionnaire. Of the 650 questionnaires distributed, 521 were returned, or 80 percent.

Among the major findings were: buidling administrators expressed a significantly more positive feeling that the clinical supervision objectives are being met than teachers. Among the perceived strengths of the program are increased teacher awareness and understanding of their teaching behavior, and movement toward successful teaching strategies as a result of the teacher/principal conferences. Although the majority of teachers perceived the clinical supervision system as supportive, they felt a need for more specific assistance. Several weaknesses int he program were perceived, including a lack of increase in trust,

rapport, and communication between teachers and adminstrators.

The purpose of the study by Akers[1] was to develop and field test an evaluation model capable of assessing secondary school facilities in West Virginia to determine compliance with standards established in the Master Plan for public education for secondary school facilities.

The Model for Secondary School Facilities in West Virginia was criteria referenced using standards established in the Master Plan for Public Education. The model was field tested in three county school districts. During the field testing process each faculty was evaluated twice. One evaluation was completed by the school principal while the second evaluation was conducted by the researcher. An in-service model was designed to standardize the application process evaluators would follow when using the Evaluation Model for Secondary School Facilities in West Virginia and all local evaluators to gather information that would permit the refinement

---

[1] Stephen Blaine Akers, "An Evaluation Model for Secondary School Facilities in West Virginia Developed in Response to the Pauley V. Bailey Decision and the Master Plan for Public Education" (Ph.D. dissertation, Virginia State University, 1984).

of both the in-service model and the Evaluation Model. The use of the Evaluation Model for Secondary School Facilities in West Virginia will provide educators with information that will assist in the development of statewide priorities for public school facilities.

Greene's[1] study attempts to build a conceptual framework focusing on the principal's role in the evaluation of teachers. To this end, a conceptual framework was developed to aid in defining and clarifying the principal's involvement in the evaluation process. It also provides a means for systematically analyzing and viewing the training or retraining of principals in the area of teacher evaluation.

The methodology utilized in the study, that a conceptual framework building, is a less finite and less restrictive in nature, design, and elements than a conceptual model. Viewed heuristically, the framework is meant to serve as ameans of discovering and thus encouraging further thought through ongoing refinement and analysis. Teacher evaluation and the principal's involvement are viewed as dynamic processes in which theory and practice are integrated through action and reflection as part of a larger

---

[1] Ibid., p. 32.

interpretive endeavor.

The framework consists of three conceptual arenas and accompanying elements. They are: (1) things over which the principal has greatest influence related to the principals' basic values, attitudes, knowledge, and skills; (2) those influences exerted by internal forces such as the superintendent, other principals, teachers and students; (3) forces farthest removed from the principal that have profound influence on how evaluation is structured. This study concluded with guidelines and discussion related to implementing the framework and recommendations for further study.

## Evaluation of Teacher Competency

In order to develop effective evaluation instruments, researchers must be cognizant of the characteristics of effective teachers; therefore, the purpose of Hague's[1] study was to compare and evaluate data on a most successful group and a least successful group of elementary teachers in order to identify attributes or characteristics which singlely or in combination were related to perceived teacher effectiveness.

---

[1] Veretta June Johnson Hague, "An Analysis of Teacher Competencies and Perceived Teacher Effectiveness (Ph.D. dissertation, Texas Tech University, 1982).

The study involved (a) attributes of age, sex, teaching experience, grade level, faculty size, school type, highest degree, certificate type, and certificate date and (b) the following Summative Teacher Evaluation (STEI) subscores: Teaching Process Sbuscore (TPS); Interpersonal Relationship Subscore (IRS); Professional Responsibilities Subscore (PRS): and Personal Characteristics Subscore (PCS). Two stratified, random samples of 50 subjects were selected from the highest (HQT) and lowest (LQT) quartiles of a group containing 700 teachers. The sample was drawn from 35 elementary schools located in a city of 174,000.

A t-test verified there were no significant differences among the variables. The Pearson Product-Moment Correlation established that significant relationships existed among variables. A chi-square technique revealed that the variable highest degree was significant.

Gaffney[1] examined competency attainment in prospective home economics teachers at Oklahoma State University in implementing instructional plans. In pursuance of the research, assessment measures were

---

[1] Betty J. Gaffney, "Assessing Competency Attainment in Prospective Home Economics Teachers in Implementing Instructional Plans" (Ph.D. dissertation, Oklahoma State University, 1975).

developed and evaluated. Three research questions were posed: (1) Do home economics education student teachers exhibit (a) competency in instructional planning, (b) specified competencies in implementing instructional plans? (2) Do secondary students gain in knowledge in classes taught by home economics student teachers who exchibit specified competencies in implementing instructional plans? (3) Is there a relationship between home economics education student teachers' exhibition of specified competencies and secondary students' gain in knowledge?

A 10-item measure was used by cooperating teachers and researcher for rating plans; a 26-item measure was used by cooperating teachers, student teachers, and researcher to assess performance; and pre- and post-tests for determining secondary student learning.

Analysis of variance by items over three rates' assessments were performed revealing that assessments on any of the measures for a given student teacher were more similar across sessions than across raters. Means scores for student teachers on the four factors were utilized in further analysis.

Among the major findings were: that student teachers, as a whole, exhibited the specified competencies in implementing instructional plans. The

secondary students learned in classes in which student teachers implemented instructional plans; that there were no significant relationships between secondary student learning and student teacher's planning and performance; and that assessment measures had applicability for assessing competency attainment.

Lydia[1] sought to determine if the process of evaluating tenure teachers made a significant difference in the academic achievement of students as determined by post-test results on the Comprehensive Test of Basic Skills (CTBS) in reading, math and language.

The CTBS in reading, math and language was administered by thirty-nine sixth grade teachers in Denver, Colorado public schools to their respective students in the Fall of 1983 pre-test and Spring of 1984 post-test. Of the total number of teachers (3() administering the test, only six were on tenure teacher evaluation. These six teachers were randomly selected from the remaining thirty-three teachers.

Multiple linear regression was used to test the following research hypothesis: the CTBS' post-test

---

[1] Johnny Lydia, "The Relationship Between Teacher Evaluation and Student Academic Achievement" (Ph.D. dissertation, University of Northern Colorado, 1984).

scores of students of teachers being evaluated are higher than the CTBA post-test scores of students of teachers not being evaluated.

The research hypothesis was rejected. There was no statistically significant difference at the .05 level between the CTBS scores of students of teachers being evaluated and the CTBS scores of students of teachers not being evaluated. The Denver public schools evaluation process of tenured teachers was designed to increase student achievement and to help teachers improve their instructional skills. The findings of this study indicated that the instructional skills of teachers being evaluated were not necessarily improved as a result of being evaluated because there was no significant difference between the achievement of their respective students and the achievement of students of teachers not being evaluated.

Gips[1] designed his study to determine whether (1) teachers can collect data on verbal interaction in the classroom by means of the Teacher College Sills and Strategies Interaction Analysis System, (2) they can code their own verbal interactions accurately, and (3) the accuracy of self-coding isdependent upon certain personality variables and upon the accuracy of the performance of the teaching models. The

hypotheses proposed were as follows: (a) that the teachers would code teacher-student verbal interactions at a mean level of at least 80 percent accuracy, (b) that the teachers would code their own verbal interactions at 80 percent or greater accuracy, (c) that there would be a positive correlation between the teacher's ability to code their own verbal interactions and their ability to perform the teaching models, and (d) that there would be a negative correlation between the accuracy of the teacher's coding of their own verbal interactions and their scores on the Dogmatism Scale.

Data for evaluation were provided through the coding and teaching performance of thirteen teachers who were trained in the ICSS Interaction Analysis System and the performance of two models of teaching, the Inductive Model and the Classroom Meeting Model. Each teacher submitted a sample of each model, and then, using the ICSS Interaction Analysis System, each teacher coded the 50 verbal interactions from each sample. The model performances were rated for their accuracy, and each participant completed the Rokeach

---

[1] Crystal Gips, "Self Analysis of Teaching Behavior: Implications for Supervision" (Ph.D. dissertation, Boston University School of Education, 1977).

Dogmatism Scale.

The data indicated that in some situations teachers can assume the responsibility for collecting information on their own teaching acts, and that some teachers are more accurate performers of this task than others. The evidence suggests some value in training teachers to be recorders of their own behavior prior to any attempts to change teaching behavior. It seems that those who were able to learn the coding system best were also willing to learn and practice new teaching behaviors.

Martin[1] assessed the effects of a self-evaluation model utilizing split-screen videotape feedback with student teachers majoring in elementary education at Bowling Green State University.

The data were collected from an enrollment population consisting of 35 student teachers registered for student teaching, Winter quarter, 1977, in the Bowling Green city schools, Eighteen participants were randomly assigned to the experimental group and twelve to the control group.

---

[1] JoAnne J. Martin, "A Study of th Effects of a Self-Evaluation Model on Focus Reaction of Student Teachers During Split-Screen Videotape Feedback" (Ph.D. dissertation, Bowling Green State University, 1977).

The sample (N=30) included student teachers in grades K-6. Subjects were videotaped and pre-tested with videotape Self-Report Form, and the videotape Observation Forms after viewing their first videotaped classroom performances.

The data were analyzed using one-way analysis of variance and analysis of co-variance with statistical significance set at the .05 level.

Among the major findings were: the experimental and control groups were not significantly different in their mean focus-reaction scores for self-teaching, and student factors on the VSRF pretest, with both groups identifying self as the factor rating most attention while students rated least attention. A significant difference was found in means focus-reaction score for self ont he VSRF post-test between subjects with experimental group students shifting focus away from self significantly more than control group students. These results indicate that self-evaluation model in conjunction with split-screen videotape feedback was effective in this study in directing subject's attention away from self characteristics toward other aspects of teaching learning situation.

## Factors Which Influence The Principal's Evaluation of Teachers

The principal problem in Skube's[1] study was to determine the effects of an in-service program of principals on the evaluation of teachers. A total of 48 principals rated 301 teachers prior to and follwoing a district's in-service program on teacher evaluation. Teachers were also in-serviced concerning self-evaluation.

The findings were as follow: (1) following the in-service, the principals who completed the teacher's ratings, rated the teachers essentially the same or not significantly lower; (2) following the in-service, the principals who completed the teacher's ratings without knowledge of the results of the teacher's self-ratings rated the teachers significantly lower in all categories than they did prior to the in-service; (3) following the in-service, the principal's and teacher's use of the rating scale indicated that the rating of teachers was more evenly distributed; (4) the effect of the rating on principal's part in-service ratings of teachers was that in all categories the mean ratings were lower when the principals rated

---

[1] Michael F. Skube, "The Effects of Principal In-Service Training on Evaluation of Elementary School Teachers" (Ph.D. dissertation, Illinois State University, 1979).

without knowledge of the teacher's self-ratings; (5) following the in-service, the teachers rated themselves significantly lower in three of the six categoreis; (6) following the in-service, the principal's ratings of teachers and the teacher's rating of themselves were more congruent; (7) the principals' in-service program in conjunction with the conditions that principals rated without knowledge of the teacher's self-ratings does effect a lowering of ratings on some of all evaluation categories and provides a more even distribution of ratings; (8) the principals' and teacher's in-service training programs were associated with essentially effecting more congruency in the principal's ratings of teachers and the teacher's ratings of themselves; (9) there is some evidence to conclude that teacher's in-service program in evaluation techniques may effect even distribution of ratings and a lowering of teacher's self-ratings in certain categories.

Whereas Skube attempted to examine the effects of in-service on teacher evaluation, LeCrone[1] endeavored to study to determine if cooperative-

---

[1] Charles Mauzey LeCrone, "The Relation of Teacher and Supervisor Competitive-Cooperative Attitudes and Principal Ratings of Successful Supervision" (Ph.D. dissertation, The University of Illinois, 1983).

competitive attitudes affected the supervisory relationship. Two unit districts in a suburban midwestern metropolitan area volunteered to participate in the study. Each supplied 60 teachers (30 elementary and 30 secondary) and their supervisors.

Three instruments were used to gather data. A cooperative-competitive Likert-type measure, the Tuckman Teacher-Supervisor Feedbak Form and a teacher-supervisor success rating form were completed by the 120 dyads.

Four hypotheses were constructed. The first three were designed to measure the effect of the cooperativeness or competitiveness by total sample, by grade level, and by school district on supervision. The fourth sought the proporation of successful supervisory relationships from among cooperative and competitive dyads.

Results from t-test and F-test scores showed no significant relationships from the first three hypotheses. Exploratory data analysis disclosed no competitive dyads; however, highly sifnificant chi-square scores were found among cooperative dyads by grade level and by school district.

McDermott's[1] study sought to examine whether supervisors whose learning styles were different would rate a teacher taught lesson differently.

Data was gathered in seven sessions from 79 individuals with regular responsibility for evaluating teachers. The procedures at the sessions involved completing of demographic data and the Transaction Ability Inventory, viewing a videotaped lesson segment, and rating the lessons on the Teacher Performance Scale. Analysis of the learning style questions and the other factors was done using either an ANOVA for differences in mean ratings of independent groups or a t-test. Variance was examined using the Bartlett test for homogeniety of variance.

The supervisors and/or evaluators learning styles were found to be related to the ratings they gave the teacher taught lesson segment. Females rated the female teacher higher than did males. No hypothesis was formulated to examine the difference in ratings from one session to another. However, evaluators from District A rated the teacher significantly lower than did evaluators from District

---

[1] James Robert McDermott, "The Learning Style of Evaluators as a Factor in the Ratings Given to a Teacher Tausht Lesson Segment" (Ph.D. dissertation, The University of Colorado, 1983).

In his study,[1] Nobles compared master's degreed teachers and bachelor's degreed teachers. The study involved 50 tenth and eleventh grade teachers of English, their students, and their principals. Half of the teachers held bachelor's degrees; half held master's degrees. Each teacher completed two instruments designed to measure her self-evaluation on selected teacher characteristics. For each teacher, one class of students completed two instruments designed to measure teacher evaluation. Each teacher's principal rated the teacher using a semantic differential.

It was found that principals rated master's degreed teachers significantly higher than did students on the following scales on the semantic differential: Fair-Unfair; Dependable-Unreliable; Sincere-Insincere; Competent-Incompetent; and Concerned-Indifferent. For each of these scales the teachers also rated themselves more highly than did their students. For the POSR master's degreed teachers rated themselves more highly on the Strict Control scale than did their students.

---

[1] Arthur Howard Nobles, "Self, tudent, and Principal Evaluation of Teachers Holding Master's Degrees and Bachelor's Degrees" (Ph.D. dissertation, The University of Arkansas, 1974).

For bachelor's degreed teachers there were significant differences between the three sources of evaluation on two scales of the semantic differential: Sincere-Insincere and Concerned-Indifferent. For both of these scales, the teachers rated themselves more highly than did their students. For the Sincere-Insincere scale the principals rated the bachelor's degreed teachers more highly than did the students. For the Concerned-Indifferent scale, the teachers rated themselves more highly than did their principal. For the POSR, no significant differences were found between teacher and student ratings.

## Summary

The review of the literature presented in this chapter was conducted in order to provide a rationale for the study. The first section of the literature review focused on various processes of teacher evaluation and the second section presented empirical studies which have been conducted on teacher evaluations. From these reviews, it become apparent that educational administrators need evaluation techniques which will help teachers to become superior instructors and thereby improve student achievement. Thus, this study attempted to determine whether one model of evaluation, the goal-setting model, was

superior to another model of evaluation, the traditional model, based on the perceptions of teachers in conjunction with student achievement.

# CHAPTER III

# METHODOLOGY

Chapter 3 discusses the methods, materials and procedures which were utilized in the study. The first section includes a concise description of the population and sample for the study. The instruments are in section two; an explanation of the methods which were utilized to collect and process the data, the research design and the analytical process are presented in section three.

## Population and Sample

The population for this study consisted of the teachers employed in the four middle schools in Area III of a large metropolitan public school system; due to their assessibility, two of the four middle schools were utilized in this study. One of the schools had an enrollment of six hundred and fifty students and sixty teachers; however, the other school had five hundred and fifty students and fifty-one teachers. Sixteen teachers from each school were randomly selected to participate in the study. In addition, the researcher obtained class rosters from each of the selected teachers. Next, the names of fifteen of the students (n=480) recorded on the roster for each of the selected teachers were

randomly selected for the teacher to obtain pre- and post-test scores for these students on the California Achievement Tests.

## Instrumentation

This study utilized the following instruments:

1. Teacher Evaluation Questionnaire
2. California Achievement Test (CAT)

Description of the instruments are provided in the succeeding sections.

### Teacher Evaluation Questionnaire (TEQ)

The Teacher Evaluation Questionnaire was constructed by Karl T. Brungarot[1] in 1979. This instrument contains thirteen items. The thirteen items are divided into four components. The components and the number of items associated with each area are shown below:

| | | |
|---|---|---|
| I. | Self Appraisal | Item 1 |
| II. | Expectation for Teacher Performance | Items 2-4 |
| III. | Evaluation Areas | Items 5-7 |
| IV. | Communication between Supervisor and Teacher | Items 8-13 |

---

[1] Karl E. Brungart, "A Procedural Model for the Improvement of Instruction through the Evaluation/Supervision of Teachers." (Ph.D. dissertation, Universtiy of Minnesota, 1979).

Subjects indicated their perceptions of each item by checking one of four alternatives - strongly agree, agree, disagree and strongly disagree. Each item was scored on a scale from 3-0. Consequently, the range of scores was from 0-39. The estimated reliability and validity for the instrument are .92 and .83 respectively.

California Achievement Tests

The California Achievement Tests were revised by Taylor, Frackenpohl and White[1] in 1969. The California Achievement tests, form C, level 17 were used in the experiment. These tests are a series of test batteries that represent a new concept in achievement testing. CAT/C and CAT/D combine the most important and useful characteristics of norm-referenced and criteria-referenced tests. This combination provides information relative to the ranking of an individual student against a norm group. It also provides specific information about the instructional needs of the student.

Test I contains 30 items that are divided into three sections. The first section requires that

---

[1] Standord E. Taylor, Helen Frackenpohl and Catherine E. White, "A Revised Core Vocabulary." Research and Information Bulletin Number 5 (N.Y. Educational Developmental Laboratories, Inc., 1969).

students identify the same meaning for an underlined word in a phrase. The second section requires that students identify the opposite meaning for an underlined word in a phase. The third section requires that students determine the meaning of multimeaning words by using sentence context.

Test II contains 40 items based on several reading selections. The test requires that students read a variety of material. This material may include biographical sketches, factual accounts, fictional passages, poems, speeches, advertisements, letters, editorials, commercials, or movie reviews. The items measure skills in literal, interpretive, and critical comprehension.

Test III contains 20 items that require students to determine which one of the underlined words in a sentence is misspelled. If all the words are spelled correctly, the students choose "none." The words included in each level of Test III reflect common spelling errors.

Language mechanics, Test IV, contains 25 items. The first 10 items measure capitalization skills. The last 15 items measure punctuation skills. The capitalization section requires that students select the part of a sentence, if any, that needs a capital letter. The punctuation section requires students to

choose whether or not one of the given punctuation marks has been left out of a sentence. All items in this test are based on rules of written standard English.

Text V contains 38 items that measure specific skills in language usage, sentence structure, and paragraph organization. The test requires students to do a variety of items that measure various skills related to effective written expressions. The items include choosing the word or words that best complete a sentence; identifying the subject or verb of a sentence; recognizing sentences that are complete, incomplete, or run-on; recognizing a sentence that is most clearly expressed; determining the best order for a group of four sentences; selecting sentences that best develop a topic sentence; or selecting the best concluding sentence. All items in this test are based on rules of written standard English.

Test VI contains 40 items that measure addition, subtraction, multiplication, and division skills. Each operation is measured by 10 items. Depending on the level of the test, the items may include operations with whole numbers, fractions, decimals, integers, or algebraic expressions.

The 45 items in Test VII measures specific

skills in understanding and using mathematics concepts. The test requires that students recognize concepts and apply problem-solving operations in various contexts. The content include numeration, number theory, number sentences, number properties, common scales, geometry, measurement, functions, graphs, or story problems.

## Collection and Processing of Data

In December 1985, the researcher contacted the assistant superintendent of a large metropolitan public school system in order to request permission to conduct the study in his school district. Once permission was granted by the assistant superintendent, permission to conduct the study at two middle schools was requested by letter from the building principals.

In the letter to the building principals (Appendix A) the researcher explained the purpose of the study and assured anonymity of the schools, teachers and students.

After permission was granted by the building principals, the researcher visited the selected schools, designated as School A and School B to: (1) randomly select sixteen of the teachers at each

school to participate in the study; (2) randomly select fifteen of the students from each of the teacher's roster in order for the teachers to obtain pre-test and post-test scores on the California Achievement tests; (3) explain the purpose of the study to the teachers; and (4) explain the goal-setting model and its administration to the principal of the school which was designated as School B in order for him to conduct the experiment at his school. The researcher conducted the experiment at School A.

On February 3, teachers at each school were randomly assigned to an experimental group (n=8) and a control group (n=8) by the researcher. The teachers in the control group were evaluated with the traditional model of evaluation as outlined in Chapter I, page 4. Moreover, the teachers in the experimental groups were evaluated with the goal-setting model as outlined in Chapter I, page 10. Teachers in the experimental and control groups administered the achievement tests during the first week and during the last week of the sixth week experiment to their selected students and recorded the results, and each teacher was pretested and posttested on the Teacher Evaluation Questionnaire to determine their perceptions of the goal-setting and and traditional models.

The researcher personally collected all questionnaires from the teachers two days after the experiment ended. Data from the instruments were analyzed with the Statistical Package for the Social Sciences,[1] which is an integrated system, especially designed for social science, including educational research.

## Research Design and Data Analysis

This study was on experimental investigation. The experiment was conducted for a period of six weeks, using the Randomized Group, Pre-test, Post-test Design.[2] When this design is employed, the dependent variable is measured before the independent variable is applied or withdrawn. Afterwards, the amount of

---

[1] Norman H. Nie, et. al., Statistical Packages for the Social Sciences, 2nd edition (New York: McGraw-Hill, 1975).

[2] Donald Ary, Lucy Jacobs and Asghar Razavieh, Introduction to Research in Education (New York: Holt, Rinehart and Winston, Inc. 1980), p. 137.

change, if any, that has taken place is computed.

The questions formulated for this study were:

1. Is there a significant difference between the teachers' perceptions of the goal-setting model and the traditional model in conjunction with student achievement?

2. Do teachers perceive the goal-setting model to be significantly different from the traditional model in:

    a. correcting weaknesses and enhancing strengths of teachers.

    b. establishing a positive working relations between the evaluatee and the evaluator?

    c. emphasizing the professional growth and needs of the teachers?

    d. focusing on the self-evaluation of the teacher?

    e. integrating individual performance objectives with the goals and objectives of the school?

    f. improving student achievement?

The t-test for independent groups was used to answer the aforementioned questions. The t-test for independent samples provides an index which is used to find the significance of the difference between the means of two groups. The groups were referred to as independent because they were drawn independently

from a population without any pairing or other relationship between the two groups.

The formula for the t-test is:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\left(\sum x_1^2 + \sum x_2^2\right)}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Where

$\bar{X}_1$ = mean for Group I

$\bar{X}_2$ = mean for Group II

$\sum X^2$ = The sum of the squared deviation score in Group I

$\sum X^2$ = The sum of the squared deviation score in Group II

$N_1$ = Number of cases in Group I

$N_2$ = Number of cases in Group II

$\bar{X}_1 - \bar{X}_2$ = Observed difference between two means

---

[1] Donald Ary, Introduction to Research in Education, p. 137.

# CHAPTER IV

## ANALYSIS OF DATA

The aim of this study was to determine whether the goal-setting model, based on the perceptions of teachers, was superior to the traditional model. In addition, it endeavored to determine whether student achievement was improved when teachers were evaluated with the goal-setting model.

This chapter is divided into two sections. Section I gives a concise description of the instruments utilized in the study, and the second section presents the results of the statistical analyses utilized to answer the research questions in Chapter I.

### Summary of Instruments

### Teacher Evaluation Questionnaire

Data pertinent to the goal-setting and traditional models were ascertained with the Teacher Evaluation Questionnaire. This questionnaire consists of thirteen questions which are divided into four components: Self-Appraisal, Expectation for Teacher Performance, Evaluation Area and Communication Between Supervisor and Teacher. Each item was scored on a scale from 3-0. Consequently, the range of scores was from 0-39 for

individual teachers.

## California Achievement Test

Student achievement was measured with the California Achievement Test. This test is divided into two categories; language arts, which has five components, and mathematics, which has two components. Raw scores were used for analytical purposes.

## Statistical Analysis

T-tests were utilized to analyze the data collected for this study. According to Ary,[1] the t-test provides an index which is used to find the significance of the difference between the means of two groups. The results of the t-tests along with the research questions are presented in this section.

1. Results with respect to question 1: Is there a significant difference between the teachers' perceptions of the goal-setting model and the traditional model and each in conjunction with student achievement?

In order to determine whether the teachers perceived the goal-setting model to be significantly different from the traditional model,

---

[1] Ary, Introduction to Research in Education, p. 21.

several t-tests were conducted on the data collected from the Teacher Evaluation Questionnaire before and after the experiment.

Table I shows the pretest and the posttest results of the t-test conducted on the perception scores of the control group, which was evaluated with the traditional model. As shown in the table, there was no significant difference between the control group's pre and posttest scores on their perceptions of the value of the evaluation model applied to them ($p>.05$, 30 df, $t=0.26$). The pre test and posttest mean scores for the control group were 20.44 and 19.76, respectively.

TABLE I

T-Test on Pretest and Posttest Data for Control Group*

| Variable | Number of Cases | Mean | Standard Deviation | Standard Error | T Value | Degrees of Freedom |
|---|---|---|---|---|---|---|
| *QTotal Group I (Pretest) | 16 | 20.44 | 8.06 | 2.01 | | |
| | | | | | 0.26 | 30 |
| Group I (Posttest) | 16 | 19.75 | 6.71 | 1.68 | | |

*Control Group is designated as Group 1
*QTotal - Total Score on Teacher Evaluation Questionnaire.

TABLE II

T-Test on Pretest and Posttest for Experimental Group*

| Variance | Number of Cases | Mean | Standard Deviation | Standard Error | T Value | Degrees of Freedom |
|---|---|---|---|---|---|---|
| QTotal Group 2 (Pretest) | 16 | 18.56 | 7.54 | 1.88 | 3.50 | 30 |
| Group 2 (Posttest) | 16 | 28.06 | 7.83 | 1.96 | | |

*QTotal - Total Score on Teacher Evaluation Questionnaire
Experimental Group - Group 2

Table II shows the pre and posttest results of the t-test conducted on the perception scores of the experimental group, which was evaluated with the goal-setting model. As can be observed in the table, there was a significant difference between the pre and posttest results for the experimental group ($p<.05$, 30df, $t=3.50$). Apparently, the t-test was found to be significant in this case because prior to the experiment, the teachers in the experimental group were responding to the traditional model, and they were adverse to it; however, they were favorable toward the goal-setting model. The mean scores as indicated in Table II were 18.56 for the pre-test and 28.06 for

the posttest. As stated earlier, this difference was significant when analyzed with the t-test.

TABLE III

T-Test on Pretest Data for Control* vs. Experimental Group*

| Variable | Number of Cases | Mean | Standard Deviation | Standard Error | T Value | Degrees of Freedom |
|---|---|---|---|---|---|---|
| *QTotal | | | | | | |
| Group 1 (Pretest) | 16 | 20.44 | 8.06 | 2.01 | 0.68 | 30 |
| Group 2 (Pretest) | 16 | 18.60 | 7.54 | 1.88 | | |

*Control Group - Group 1
*Expremental Group - Group 2
*QTotal - Total Score on Teacher Evaluation Questionnaire

Table 3 shows the pretest results of the t-test conducted on the control group versus the experimental group. The table shows that the mean score for the control group was 20.44; moreover, the mean score for the experimental group was 18.60. When a t-test was applied to these two means no significant difference was found to exist between them ($p>.05$, 30 df, $t=0.68$). It would appear, then, that prior to the administration of

the treatment to the experimental group, the control and experimental groups had similar feelings about evaluation models, and since the mean scores were rather low, it would appear that the teachers had an adversion to the evaluation models. It was interesting to note, however, that after treatment had been administered to the experimental group (see Table IV), the mean score for this group increased from 18.60 to 28.06; however, the mean score for the control group decreased from 20.44 to 19.00. The t-test which was applied to these two means showed that there was a significant difference between them ($p<.05$, 30 df, $t=3.23$). These results show that the teachers in the experimental group were more favorable toward the goal-setting model than the teachers in the control group were toward the traditional model. It would appear, then, that based on the perceptions of teachers, the goal-setting model is superior to the traditional model of evaluation.

Several pre and post t-tests were also conducted in order to determine the effects of the goal-setting model and the traditional model on student achievement. Table V indicates that the pre and posttest results in language arts for the

TABLE IV

T-Test on Posttest Data for Control* versus Experimental Group

| Variable | Number of Cases | Mean | Standard Deviation | Standard Error | T Value | Degrees of Freedom |
|---|---|---|---|---|---|---|
| *QTotal | | | | | | |
| Group 2 | 16 | 28.06 | 7.83 | 1.96 | 3.23 | 30 |
| Group 1 | 16 | 19.75 | 6.71 | 1.68 | | |

TABLE V

T-Test on Posttest Data for Control Group on TSCATLAN*

| Variable | Number of Cases | Mean | Standard Deviation | Standard Error | T Value | Degrees of Freedom |
|---|---|---|---|---|---|---|
| *TSCATIAN | | | | | | |
| Group 1 Traditional (Pretest) | 16 | 850.69 | 885.83 | 221.46 | 0.19 | 30 |
| Group 1 Traditional (Posttest) | 16 | 913.19 | 949.14 | 237.29 | | |

*QTotal - Total Score on Teacher Evaluation Questionnaire
*Control Group is Group 1
*Experimental Group is Group 2
*TSCATLAN - Total Score on Language Arts Section of CAT

to the evaluation models. It was interesting to note, however that after treatment had been administered to the experimental group (See Table 4), the mean score for this group increased from 18.60 to 28.06; however, the mean score for the control group decreased from 20.44 to 19.00. The t-test which was applied to these two means showed that there was a significant difference between them ($p<.05$, 30 df, $t=3.23$). These results show that the teachers in the experimental group were more favorable toward the goal-setting model than the teachers in the control group were toward the traditional model. It would appear, then, that based on the perceptions of teachers, the goal-setting model is superior to the traditional model of evaluation.

Several pre and post t-tests were also conducted in order to determine the effects of the goal-setting model and the traditional model on student achievement. Table V indicates that the pre and posttest results in langauge arts for the control and/or traditional group were not significant ($p>.05$, 30 df, $t=0.19$). The pre and post mean scores for the control group in language arts were 850.69 and 913.19, respectively. Similarly, the t test applied to the pre and posttest means (927.94 and 1162.75 respectively) in Table VI for the experimental and/or goal-setting roup in language arts was not significantly different ($p>.05$, 30 df, $t=0.61$).

control and/or traditional group were not significant ($p>.05$, 30 df, $t=0.19$). The pre and post mean scores for the control group in language arts were 850.69 and 913.19, respectively. Similarly, the t-test applied to the pre and posttest means (927.94 and 1162.75 respectively) in Table VI for the experimental and/or goal-setting group in language arts was not significantly different ($p>.05$, 30 df, $t=0.61$).

TABLE VI

T-Test on Pre and Posttest Data for Experimental Group*

| Variable | Number of Cases | Mean | Standard Deviation | Standard Error | T Value | Degrees of Freedom |
|---|---|---|---|---|---|---|
| *TSCATLAN | | | | | | |
| Group 2 (Pretest) | 16 | 927.94 | 966.62 | 241.66 | 0.61 | 30 |
| Group 2 (Posttest) | 16 | 1162.75 | 1204.60 | 301.15 | | |

TABLE VII

T-Test on Pretest Data for Control versus Experimental Group

| Variable | Number of Cases | Mean | Standard Deviation | Standard Error | T Value | Degrees of Freedom |
|---|---|---|---|---|---|---|
| TSCATLAN | | | | | | |
| Group 1 | 16 | 850.69 | 885.83 | 221.46 | 0.24 | 30 |
| Group 2 | 16 | 927.94 | 966.62 | 241.66 | | |

*Control Group - Group 1
*Experimental Group - Group 2
*TSCATLAN - Total Score on Language Arts

TABLE VIII

T-Test on Posttest Data on Control* versus Experimental Group*

| Variable | Number of Cases | Mean | Standard Deviation | Standard Error | T Value | Degrees of Freedom |
|---|---|---|---|---|---|---|
| *TSCATLAN | | | | | | |
| Group 1 | 16 | 913.19 | 949.14 | 237.29 | 0.65 | 30 |
| Group 2 | 16 | 1162.75 | 1204.60 | 301.15 | | |

Table VII shows that the pretest results for the control group versus the experimental group with mean scores of 850.69 and 917.94, respectively, in language arts were not significantly different ($p>.05$, 30 df, $t=.024$). Likewise Table VIII indicates that the posttest results for the control versus the experimental group with means of 913.19 and 1162.75 in language arts, respectively, were not significantly different ($P>.05$, 30 df, $t=0.65$). It would seem to appear, then, that the type of model used to evaluate teachers does not influence student achievement in language arts. It was, however, interesting to note that the students of teachers who were evaluated with the goal-setting model showed more improvement in the post mean score than the students of teachers who were evaluated with the traditional model (compare mean scores in Tables VII and VIII).

The data in Table XIX indicate that the pre and posttest results for the control group in mathematics was not significant ($p>.05$, 30 df, $t=.01$). The mean scores were 480 for the control group, and 479 for the experimental group. Similarly, Table X shows that the pre and posttest

results for the experimental group in mathematics were not significantly different ($p > .05$, 30 df, $t = .33$). The mean score for the pretest was 444, and the mean score for the posttest was 502.

TABLE XIX

T-Test on Pre and Postest Data for Control Group

| Variable | Number of Cases | Mean | Standard Deviation | Standard Error | T Value | Degrees of Freedom |
|---|---|---|---|---|---|---|
| **TSCATMAT Group 1 (Pretest) | 16 | 480 | 496.41 | 124 | 0.01 | 30 |
| Group 1 (Posttest) | 16 | 479 | 494.53 | 123 | | |

*Control Group - Group 1
*Experimental Group - Group 2
*TSCATLAN - Total Score on Language Arts
*TSCATMAT - Total Score in Mathematics

TABLE X

T-Test on Pretest and Posttest Data for Experimental Group*

| Variable | Number of Cases | Mean | Standard Deviation | Standard Error | T Value | Degrees of Freedom |
|---|---|---|---|---|---|---|
| *TSCATMAT Group 2 (Pretest) | 16 | 443.81 | 466.76 | 116.69 | 0.33 | 30 |
| Group 2 (Posttest) | 16 | 501.69 | 518.17 | 129.54 | | |

*Experimental Group - Group 2
*TSCATMAT - Total Score on Mathematics

TABLE XI

T-Test on Posttest Data for Control versus Experimental Group

| Variable | Number of Cases | Mean | Standard Deviation | Standard Error | T Value | Degrees of Freedom |
|---|---|---|---|---|---|---|
| *TSCATMAT Group 1 | 16 | 480 | 496 | 124 | | |
| Group 2 | 16 | 444 | 467 | 117 | 0.21 | 30 |

TABLE XII

T-Test on Posttest Data for Control versus Experimental Group

| Variable | Number of Cases | Mean | Standard Deviation | Standard Error | T Value | Degrees of Freedom |
|---|---|---|---|---|---|---|
| TSCATMAT | | | | | | |
| Group 1 | 16 | 479 | 495 | 124 | 0.13 | 30 |
| Group 2 | 16 | 502 | 518 | 130 | | |

*Control Group - Group 1
*Experimental Group - Group 2
*TSCATMAT - Total Math Score

Tables 11 and 12 indicate that the pre and posttest results for the control versus the experimental groups in mathematics were not significantly different ($p > .05$, 30 df, $t=.21$, for the pretest, and $p>.05$, 30 df, $t=.13$ for the posttest). The pretest mean scores were 480 for the control group and 444 for the experimental group. The posttest mean scores were 479 for the control group and 502 for the experimental group. It was interesting to note that even though there was no significant difference between the pre and posttest scores for the two groups, the post mean math score for the experimental group increased from 444 to 502; moreover, the post mean math score

for the control group dropped from 480 to 478. Even though the post mean score for the experimental group increased, the increase was not great enough to be significant. Thus, the results of this study show that student achievement is not significantly influenced by the goal-setting model.

2. Results with respect to question 2a: Do teachers perceive the goal-setting model to be significantly different from the traditional model in correcting weaknesses and enhancing strengths of teachers?

TABLE XIII

T-Test on Posttest Data for Control* versus Experimental Group*

| Variable | Number of Cases | Mean | Standard Deviation | Standard Error | T Value | Degrees of Freedom |
|---|---|---|---|---|---|---|
| *QTotal | | | | | | |
| Group 2 | 16 | 28.06 | 7.83 | 1.96 | 3.23 | 30 |
| Group 1 | 16 | 19.75 | 6.71 | 1.68 | | |

*Control is Group 1.
*Experimental is Group 2.
QTotal - Total Score on Teacher Evaluation Questionnaire

As indicated in Table XIII, the teachers in the experimental group perceived the goal-setting model to be superior to the traditional model in correcting the weaknesses and enhancing strengths of teachers. In other words, the overall usefulness of the goal-setting model was superior to the traditional model. The mean score for the control group was 19.75 and the mean score for the experimental group was 28.06. The t-test applied to these two means revealed that they were significantly different ($p<.05$, 30 df, $t=3.23$). It would seem to appear, then, that the goal-setting model, which had a higher mean score than the traditional model was superior to the traditional model in correcting weaknesses and enhancing strengths of teachers, as indicated in question 1.

3. Results with respect to question 2b: Do teachers perceive the goal-setting model to be significantly different from the traditional model in establishing a positive working relationship between the evaluatee and the evaluator?

As can be observed in Table XIV, teachers perceived the goal-setting model to be superior to the traditional model in establishing a positive

working relationship between the evaluatee and evaluator. The questionnaire questions utlized to answer this question were questions five and six (See Appendix). The mean score for the control/traditional group was 2.44. A t-test applied to the means fo the two groups showed them to be significantly different ($p<.05$, 30 df, $t=43.75$).

TABLE XIV

T-Test on Question 5 and 6

| Variable | Number of Cases | Mean | Standard Deviation | Standard Error | T Value | Degrees of Freedom |
|---|---|---|---|---|---|---|
| Question 5 and 6 | | | | | | |
| Group 1 (Control) | 16 | 1.69 | 0.793 | 0.198 | 43.75 | 30 |
| Group 2 (Experi-mental) | 16 | 2.43 | 0.727 | 0.182 | | |

4. Results with respect to question 2c: Do teachers perceive the goal-setting model to be significantly different from the traditional model in emphasizing the professional growth and needs of the teachers?

Question 4 on the Teacher Evaluation Questionnaire (see Appendix) was utilized to answer this question. Table XV shows that the mean score for the experimental group was 2.50 and the mean score for the control group was 1.63. A t-test applied to the means of these two groups showed them to be significantly different ($p<.05$, 30 df, $t=11.91$). Thus, in this study, teachers perceived the goal-setting model to be superior to the traditional model in emphasizing the professional growth and needs of teachers.

TABLE XV

T-Test on Question 4

| Variable | Number of Cases | Mean | Standard Deviation | Standard Error | T Value | Degrees of Freedom |
|---|---|---|---|---|---|---|
| Question 4 | | | | | | |
| Group 1 (Control) | 16 | 1.63 | 0.806 | 0.202 | 11.91 | 30 |
| Group 2 (Experimental) | 16 | 2.50 | 0.516 | 0.129 | | |

5. Results with respect to question 2d: Do teachers perceive the goal-setting model to be significantly different from the traditional model in focusing on the self-evaluation of the teacher?

TABLE XVI

T-Test on Questions 1 and 3

| Variable | Number of Cases | Mean | Standard Deviation | Standard Error | T Value | Degrees of Freedom |
|---|---|---|---|---|---|---|
| Questions 1 and 3 | | | | | | |
| Group 1 (Control) | 16 | 1.57 | 0.892 | 0.223 | 23.79 | 30 |
| Group 2 (Experimental) | 16 | 2.25 | 0.775 | 0.194 | | |

Questions one and three on the Teacher Evaluation Questionnaire were used to answer this question. The t-test results found in Table XVI provide a graphic of the relationship discovered. As can be observed, the mean score for the control group was 1.57, and the mean score for the experimental group was 2.25. The t-test applied to the means of these two groups showed them to be significant ($p<.05$, 30 df, $t=23.79$). It appears, then, that based on the perceptions of the teachers, the goal-setting model is superior to the traditional model in focusing on the self-evaluation of teachers.

6. Results with respect to question 2e: Do

teachers percieve the goal-setting model to be significantly different from the traditional model in integrating individual performance objectives with the goals and objectives of the school?

TABLE XVII

T-Test on Questions 9 and 12

| Variable | Number of Cases | Mean | Standard Deviation | Standard Error | T Value | Degrees of Freedom |
|---|---|---|---|---|---|---|
| Questions 9 and 12 | | | | | | |
| Group 1 (Control) | 16 | 0.31 | .060 | .015 | 6.5 | 30 |
| Group 2 (Experimental) | 16 | 0.44 | .142 | .035 | | |

TABLE XVIII

T-Test on Question 13

| Variable | Number of Cases | Mean | Standard Deviation | Standard Error | T Value | Degrees of Freedom |
|---|---|---|---|---|---|---|
| Question 13 | | | | | | |
| Group 1 (Control) | 16 | 1.44 | .814 | .203 | 27 | 30 |
| Group 2 (Experimental) | 16 | 2.25 | .931 | .233 | | |

Questions nine and twelve on the Teacher Evaluation Questionnaire were utilized to answer this question. As indicated in Table XVII, the mean score for the control group was 0.31 and the mean score for the experimental group was 0.44. The t-test applied to the means showed them to be significantly different ($p<.05$, 30 df, $t=6.5$). Thus, it appears that the goal-setting model is superior to the traditional model in integrating individual performance objectives with the goals and objectives of the school.

7. Results with respect to question 2f: Do teachers perceive the goal-setting model to be significantly different from the traditional model in improving student achievement?

As shown in Table 18, the teachers in the experimental group perceived the goal-setting model to be superior to the traditional model in improving student achievement. Question 13 on the Teacher Evaluatlion Questionnaire was used to answer the research question. The mean score for the control group was 1.44 and the mean score for the experimental group was 2.25. The t-test applied to these two means revealed that they were

significantly different ($p<.05$, 30 df, $t=27$). It was interesting to note that even though the teachers perceived the goal-setting model to be superior to the traditional model in improving student achievment, when the language arts and mathematics portions of the California Achievement Test were administered to students of teachers in the experimental group, no significant difference was found to exist between the traditional model and the goal-setting model in terms of influencing student achievement. However, it was noted earlier that the post mean score did improve for students whose teachers were evaluated with the goal-setting model. It is therefore conceivable that, since this experiment only lasted 6 weeks, a similar experiment administered over a longer period of time may find the goal-setting model to be superior in improving student achievement.

## Summary

The purpose of this chapter was to present the statistical analysis of the data with respect to the traditional model of evaluation and the goal-setting model in conjunction with student achievement.

Based on the perceptions of teachers, significant relationships were found to exist

between the traditional model and the goal-setting model, with the goal-setting model being superior to the traditional model. In addition, teachers found the goal-setting model to be superior to the traditional model in the following areasd: correcting weaknesses and enhancing strengths of teachers, establishing a positive working relationship between evaluatee and evaluator, emphasizing the professional growth and needs of the teacher, integrating individual performance objectives with goals and objectives of the school, focusing on the self-evaluation of the teacher, and in improving student achievement; however, the language arts and mathematics components of the California Achievement Test did not find the goal-setting model to be superior to the traditional model in improving student achievement.

In chapter 5, the results are summarized and discussed. Conclusions and recommendations for further research are also presented.

# CHAPTER V

## SUMMARY, CONCLUSIONS, RECOMMENDATIONS

The fifth chapter of this study is divided into three sections. A summary of the purpose of the study along with the methodology and statistical procedures utilized in the study is presented in section one. The findings of the study are summarized and conclusions are presented in the second section, and recommendations for further study are outlined in the third section.

### Summary of the Study

The purpose of this study was to determine whether the goal-setting model, based on the perception of teachers, was superior to the traditional model. In addition, it endeavored to determine whether student achievement was improved when teachers were evaluated with the goal-setting model.

To fulfill the purpose of this, the following research questions were formulated:

1. Is there a significant difference between the teachers' perceptions of the goal-setting model and the traditional model in conjunction with student achievement?

2. Do teachers perceive the goal-setting to be significantly different from the traditional model in:
   a. correcting weaknesses and enhancing strengths of teachers?
   b. establishing a positive working relationship between the evaluatee and the evaluator?
   c. emphasizing the professional growth and needs of the teachers?
   d. focusing on the self-evaluation of the teacher?
   e. integrating individual performance objectives with the goals and objectives of the school?
   f. improving student achievement?

The population for this study consisted of two of the four middle schools in a large metropolitan school district. Sixteen of the teachers at each school and fifteen of the students of each teacher were randomly selected to participate in this study.

Data for the study were collected by using two instruments; the Teacher Evaluation Questionnaire, which measured the teachers' perception of the goal-

setting model and the traditional model and the California Achievement Test, which measured student achievement in language and mathematics. These instruments were administered to the teachers and students before and after the six week experiment.

The data obtained from the instruments were analyzed with the Statistical Package from the Social Science (SPSS)[1] at the Atlanta University Computer Center. T-tests were used to analyze the data.

Findings and Conclusions of the Study

The findings/results of the study are summarized in this section and pertinent conclusions are stated. The results and conclusions are presented with respect to each research question.

1. Results and conclusions with respect to question 1: Is there a significant between the teachers' perceptions of the goal-setting model and the traditional model and each of these in conjunction with student achievement.

The pretest and posttest results for the traditional group on the Teacher Evaluation Questionnaire was not significant at the 0.5 level; the t-value was 0.26; however, the pretest and

---

[1] Nie, Statistical Package for The Social Sciences.

posttest results fro the goal-setting model on the Teacher Evaluation Questionnaire were significant. The t-value was 3.50. As explained in Chapter IV, teachers wer probably more favorable about the goal-setting model after being evaluated by it, and this caused the pre and posttest results to be significantly different.

The pretest results for the traditional versus the goal-setting model were not significant; the t-value was .68; however, the posttest results for the traditional versus goal-setting model were significant, with a t-value of 3.23. Thus, this finding shows that the teachers perceived the goal-setting model to be superior to the traditional model.

The goal-setting model was probably found to be superior to the traditional model because, according to McGreal (see p. 12), it allows teachers to take an active role in the evaluation process. This is important because, according to Blake and Mouton,[1] the participative approach brings out all concerns, avoids crises, and helps subordinates to develop the competencies essential for effective

---

[1] Robert Blake and Jane Mouton. The Managerial Grid (New York: Harper and Row, 1954).

participation and more action. They stated further that objectives can be communicated more clearly when all have participated in their development. Likewise, Spears (see page 45) concluded from his study that teachers and all those concerned should be actively involved in the evaluation process because it increases competency; moreover, Willard (see page 48) concluded that teachers and supervisors prefer an evaluation model which encourages growth, which is one of the characteristics of the goal-setting model.

The goal-setting model also focuses on the needs of the teacher. This is very important because, according to Maslow,[1] people have needs and when these needs are met or if there is an attempt to meet these needs, people become more motivated. In essence, the goal-setting model was probably found to be superior to the traditional model because it has a high concern for the teacher and the school.

Even though the goal-setting model was perceived to be superior to the traditional model, it was not found to be superior to the traditional model in improving student achievement on the California Achievement Test. The following insignificant t-values were observed for student achievement; pre and

---

[1] Abraham Maslow, Motivation and Personality, (New York, Haprer and Row, 1960):6.

posttest results for the traditional group in language were .19; pre and posttest results for the goal-setting group in language were .61; pretest results for the traditional versus goal-setting in language were .24; posttest results for traditional versus goal-setting group in language were .65; pre and posttest results for the traditional group in math were .01; pretest and posttest results for the goal-setting were .33; pretest results for the traditional versus the goal-setting group in math were .21 and posttest results for the traditional versus the goal-setting group were .13.

Even though the researcher could not find any other studies dealing with teacher evaluation and student achievment it is the contention of the researcher as stated in Chapter IV that since post mean scores were higher for students whose teachers were evaluated with the goal-setting mdoel, a similar study over a longer period of time should find the goal-setting model to be superior to the traditional model. This is logical because the goal-setting model as previously stated allows teachers to take an active role in the evaluation process, and according to Getzels and Guba,[1] when people are involved they feel

---

[1] Jacob W. Getzels and Eugene G. Guba, "Social Behavior," The School Review (Winter, 1957):423-441.

more committed to an organization and are therefore more enthusiastic. It would seem to appear then, that student achievement will be higher when teachers are evaluated with the goal-setting model because due to their personal involvement teachers morale will probably be higher.

Kaura[1] found that student achievement increases with students who have teachers who are highly motivated.

2. Results and conclusions with respect to question 2a. Do teachers perceive the goal-setting model to be significantly different from the traditional model in correcting weaknesses and enhancing strengths of the teacher?

A t-test applied to the post means of the goal-setting and traditional models was significant at 3.23. As indicated earlier, this shows that the teachers feel that the goal-setting model is far more useful in correcting weaknesses and enhancing strengths of teachers than the traditional model. They therefore perceive the goal-setting model to be superior to the traditional model. Thus, as

---

[1] Hussen S. Kaura, "Student Achievement and Teacher Morale," (Ed.D. Dissertaiton, University of Michigan, 1963):94.

understandable because, as the researcher discussed earlier, the goal-setting is deeply rooted in motivation theory with an emphasis on the needs of teachers. It is only natural then, that teachers would perfer the goal-setting model. Based on this finding, it was concluded that teachers prefer an evaluation model that is teacher centered rather than one which is centered on the evaluator.

3. Results and conclusions with respect to question 3b: Do teachers perceive the goal-setting model to be significantly different from the traditional model in establishing a positive working relationship between the evaluatee and the evaluator?

A t-value of 43.75 showed that the teachers perceived the goal-settingmodel to be far more superior than the traditional model in establishing a positive working relationship between the evaluatee and the evaluator. Acccording to McGreal (see page 9) the goal-setting model fosters a positive working relationship because it allows the teachers to have imput into the development and revision of evaluation process, and it provides for contact between the teacher and the evaluator before, during, and after the evaluation.

In his discussion of the goal-setting model, Iwanick (see page 11) stated that the pre-conference that is used in the goal-setting model clearly sets it off from the traditional model, and it is the most valuable and the most important activity of this process because it allows the teacher to have some initial input, and it gives the teacher and the supervisor a chance to get to know each other. According to Zelenak (see page 6); however, the traditional model does not promote a positive working relationship between teacher and evaluator because it promotes low teacher envolvment and minimal contact time between supervisors and teachers. Based on this finding, then, it was concluded that positive relationships between teachers and supervisors will improve when teachers are allowed to actively participate in the evaluation process, and there is a familiarity between the teachers and the supervisors.

4. Results and conclusions with respect to question 2c: Do teachers perceive the goal-setting model to significantly different from the traditional model in emphasizing the personal growth of the teacher?

A t-value of 11.91 showed that the teachers perceived the goal-setting model to be superior to the

traditional model in emphasizing the growth and needs of the teachers.

In this study, the teachers perceived the goal-setting model to be superior to the traditional in emphasizing the growth and needs of the teachers. The t-value was 23.79. According to Iwinick (see page 10) the goal-setting model promotes growth because it gives the teacher a chance to identify her weaknesses and confer with the supervisor in the initial conference about strategies which will help her to overcome those weaknesses. Once a teacher knows her weaknesses and works toward eliminating them, this will help her to grow and of course become a better teacher. The traditional model, however, does not foster growth because it does not focus on the needs of the teachers according to Bolton (page 5). The traditional model promotes the use of evaluative data gathered only for administrative purposes, not for the growth/improvement of teacher performance. Based on this finding, it was concluded that an effective evaluation model must promote growth by helping teachers to identify their weaknesses and thereby help them to eliminate those weaknesses.

5. Results and conclusions with respect to question 2d: Do teachers perceive the goal-setting model to be significantly different

from the traditional model by focusing on the self evaluation of teachers?

Teachers found the goal-setting model to be significantly superior to the traditional model in focusing on the self evaluation of teachers. This is evident because according to Iwinick (page 10) one of the first steps in the goal-setting model is for teachers to conduct a self evaluation and identify areas of improvement, and based on the teachers evaluation the teacher, along with the supervisor, set goals for the self improvement. According to Oldham (page 27) it is important because it allwos the teachers' behavior to be changed by the teacher without outside control, and this is important because true behavioral change occurs only when the individual acknowledges and internalizes discrepancies in his/her behavior through his/her own violition. Based on the findings, it was concluded that an effective evaluation model should allow teachers to evaluate themselves in conjunction with their supervisors.

6. Results and conclusions with respect to question 2e: Do teachers perceive the goal-setting to be significantly different from the traditional in intergrating individual performance objectives with the goals and objectives of the school?

A t-value of 6.5 indicated that the teachers perceived the goal-setting model to be superior to the traditional model in combining individual objectives of the teacher with school objectives. This is important because as cited earlier, Getzels and Guba stated that when individual needs/goals are integrated with the goals of the organization, motivation increases. This is important for teachers because, as stated by Kaura, when teachers are motivated, student achievement seems to increase. Based on the finding of this study, it was concluded an effective evaluation model should integrate individual goals with school goals.

7. Results and conclusions with respect to question 2f: Do teachers perceive the goal-setting model to be significantly different from the traditional model in improving student achievement?

A t-value of 27 showed that teachers perceived the goal-setting model to be superior to the traditional model in improving student achievement; however, the results of the California Achievement Test did not show the goal-setting model to be superior to the traditional in improving student achievement as stated in Chapter IV, the California

Achievment Test may not have been significant because the experiment only lasted six weeks, and this was probably not enough time for significant achievment to take place.

## Summary of Findings

1) Teachers perceived the goal-setting model to be superior to the traditional model; however, the goal-setting model was not significantly related to student achievement when analyzed with the California Achievement Test.
2) Teachers perceived the goal-setting model to be superior to the traditional model in the following areas: correcting weaknesses and enhancing strengths of teachers; establishing a positive working relationship between the evaluatee and the evaluator; emphasizing the professional growth and needs of the teachers, focusing on the self-evaluation of teachers; integrating individual performance goals with goals of the school; and improving student achievement.

## Recommendations

1. It is recommended that additional studies utilizing larger samples in a more diverse

population over a longer period of time to replicate this research. Teachers and students from all grade levels should be used to expand the scope of the study.

2. Since the findings for the goal-setting and traditional models were based solely on the teachers' perceptions other studies should be conducted to determine to what extent the teacher's perceptions are congruent with more of the perceptions concerning these two models.
3. It is recommended the educational administrators familiarize themselves with the current research on effective evaluation models and utilize the research in their endeavors to improve teachers' performance.
4. It is recommended that workshops and seminars on evaluation be held simultaneously for teachers and principals.
5. It is recommended that educational administrators select an evaluation model which addresses the needs of the teachers.
6. School principals should make school goals clear to teachers so teachers may be able to integrate their own objectives to the objectives of the school.

# BIBLIOGRAPHY

# BIBLIOGRAPHY

## Books

Aleomoni, Lawrence M. Evaluation by Students to Identify Instructional Problems. Urbana, Illinois: Office of Instrtuctional Resources, Illinois University, 1973.

Ary, Donald; Jacobs, Lucy; and Razavieh, Asghar. Introduction to Research in Education. New York: Holt, Rinehart and Winston, Inc., 1980.

Blake, Robert and Mouton, Jane. The Managerial Grid, New York, Harper and Row, 1954.

Bolton, D. L. Selection and Evaluation of Teachers. Berkeley, California: McCutchan, 1973.

Crim, Alonzo. Evaluation Guide. Atlanta, Georgia, 1980.

Hyman, R. J. School Administrators Handbook of Teacher Supervision and Evaluation Methods. Englewood Cliffs, New Jersey: Prentice-Hall, 1975.

Iwanicki, E. F. Handbook of Teacher-Evaluation. Beverly Hills, California: Sage, 1981.

McGreal, Thomas. Successful Teacher Evaluation. Virginia: ASCD, 1983.

Nie, Norman H., et al. Statistical Packages for Social Science, 2d ed. New York: Holt, Rinehart and Winston, Inc., 1980.

Trovers, R. M. W. Handbook of Teacher Evaluation. Beverly Hills, California: Sage, 1981.

## Periodicals

Drumheller, Sidney J. "Evaluating Teachers Through a Jaundiced Eye." Educational Technology 14:7 (1974):17-22.

Getzels, Jacob and Guba, Eugene. "Social Behavior and the Administrative Process." The School Review 65 (Winter, 1957):423-441.

McGreal, Thomas. "Helping Teachers Set Goals." Educational Leadership 37 (February 1980):414-420.

National Education Association. "Planning the Evaluation of Education," Washington, D. C.: NEA Division of Instruction and Professional Development, 1972.

Oldham, Neil, "Evaluating Teachers for Professional Growth: Current Trends in School Policies and Programs," An Education USA Special Report. Virginia: NSSPA, 1974.

Popham, W. J. "Teacher Evaluation or the Wrong Test for the Right Job." Paper Presented at the Annual Meeting of the American Educational Research Association. Los Angeles, California, 1981.

Taylor, Stanford E.; Frankenpohl, Helen; and White, Catherine E. "A Revised Core Vocabulary." Research Bulletin Number 5. New York: Educational Developmental Laboratories, Inc., 1969.

Wagoner, R. L. and O'Hanlon, J. P. "Teacher Attitudes Toward Evaluation." Journal of Teacher Evaluation 19 (1968):471-475.

Zelenak, M. J. and Snider, B. C. "Teachers Don't Resent Evaluation If It's for Improvement of Instruction." Phi Delta Kappan 55 (1974):348-349.

## Dissertations

Adams, Gwendolyn J. "Development and Field Testing of An Evaluation Instrument Assessing Competencies of Teachers of Learning Disabled Students." Ph.D. dissertation, Auburn University, 1977.

Akers, Stephen Blaine. "An Evaluation Model for Secondary School Facilities in West Virginia Developed in Response to the Pauley Baily Decision and the Master Plan for Public Education." Ph.D. dissertation, Virginia State University, 1984.

Brungardt, Karl. "A Procedural Model for the Improvement of Instruction Through the Evaluation/Supervision of Teachers." Ph.D. dissertation, The University of Minnesota, 1979.

Gaffney, Betty J. "Assessing Competency Attainment in Prospective Home Economics Teachers in Implementing Instructional Plans." Ph.D. dissertation, Oklahoma State University, 1975.

Gips, Crystal. "Self Analysis of Teaching Behavior: Implications for Supervision." Ph.D. dissertation, Boston University School of Education, 1977.

Greene, Linda Hincher. "A Conceptual Framework for Principals' K-12 Involvement in the Evaluation of Teachers." Ph.D. dissertation, The University of North Carolina, 1984.

Hague, Veretta June Johnson. "An Analysis of Teacher Competencies and Perceived Teacher Effectiveness." Ph.D. dissertation, Texas Technical University, 1982.

Hassein, S. Kaura. "Student Achievement and Teacher Morale." Ed.D. Dissertation, University of Michigan, 1963.

Heebink, William. "A Comprehensive Investigation of a Developmental Teacher Evaluation Program's Effectiveness with Experienced Elementary Teachers in an Urban Wisconsin School District." Ph.D. dissertation, Florida State University, 1979.

Heuss, Ronald Neal. "Teacher Evaluation Purposes, Procedures, and Instruments in Texas Public Schools." Ph.D. dissertation, Baylor University, 1984.

Inglis, Sandra. "The Development and Validation of an Instrument to Assess Teacher Irritations and Teacher Effectiveness as Reported by Students in a Technical and General Post-Secondary Educational Setting." Ph.D. dissertation, The University of Florida, 1976.

Janney, Patra Moss. "Teacher Evaluation Guidelines Based on Expert Judgment, National Laws, and State Laws of West Virginia." Ph.D. dissertation, Virginia State University, 1984.

Le Crone, Charles Mauzey. "The Relation of Teacher and Supervisor Competitive-Cooperative Attitudes and Principal Ratings of Successful Supervision." Ph.D. dissertation, The University of Ilinois, 1983.

Leeth, Benton M. "The Devleopment and Validation of an Industrial Arts Teaching Rating Instrument." Ph.D. dissertation, Texas Agricultural and Mechanical University, 1977.

Lydia, Johnny. "The Relationship Between Teacher Evaluation and Student Academic Achievement." Ph.D. dissertation, University of North Colorado, 1984.

Martin, JoAnne J. "A Study of the Effects of a Self-Evaluation Model on Focus Reaction of Student Teachers During Split-Screen Videotape Feedback." Ph.D. dissertation, Bowling Green State University, 1977.

McDermott, James Robert. "The Learning Style of Evaluators as a Factor in Ratings Given to a Teacher Taught Lesson Segment." Ph.D. dissertation, The University of Colorado, 1983.

Nobles, Arthur Howard. "Self, Student, and Principal Evaluation of Teachers Holding Master's Degrees and Bachelor's Degrees." Ph.D. dissertation, The University of Arkansas, 1974.

Schmitt, Edward L. "A Cost Effectiveness Analysis of Selected Teacher Performance Evaluation Systems." Ph.D. dissertation, Iowa State University, 1976.

Scime, Joseph. "The Conflict Between Evaluation and Supervision: A Study of Clinical Supervision Program Designed to Link Instruction Supervision and Evaluation of Staff Development." Ph.D. dissertation, George Peabody Teachers College, 1984.

Schirer, William Roe. "The Purposes of School Evaluation as Perceived by Teachers, Principals, Superintendents, and School Board Presidents of Schools Holding Membership in Wisconsin North Central Association." Ph.D. dissertation, The University of Wisconsin, 1980.

Skuhe, Michael F. "The Effects of Principal in-Service Training on Evaluation of Elementary School Teachers." Ph.D. dissertation, Illinois State University, 1979.

Spears, Voy. "Guidelines for Developing a Program of Teacher Evaluation." Ph.D. dissertation, The University of Arkansas, 1974.

Swanson, Ronald George. "Teacher Evaluation as a Function of Leadership Systle: A Multiple Correlational Approach." Ph.D. dissertation, North Texas State University, 1974.

Willard, David T. "An Assessment of the Effects of a Staff Evaluation Model Developed from Douglas McGregor's Theory "Y" Upon the Attitudes fo Teachers and Supervisors Toward Evaluation." Ph.D. dissertation, St. Louis University, 1979.

Williams, Jane L. "Analysis of Teacher Evaluation Instruments in Use in Tennessee Public School Systems." Ph.D. dissertation, East Tennessee State University, 1983.

# APPENDICES

# APPENDIX A

# CORRESPONDENCE

Mrs. Anne L. Crawford
1423 Mill Acres Drive, S. W.
Atlanta, Georgia 30311

Dear Principal:

I am pursuing a doctoral degree in Educational Administration at Atlanta University. As a part of my doctoral work, I have chosen the following topic for investigation: Teachers' Perceptions of the Goal-Setting Model and the Traditional Model of Teacher Evaluation in Conjunction with Variables and Student Achievement. This study is under the direction of Dr. Ganga Persaud, Professor and Chairperson of the Department of Educational Administration and Policy Studies at Atlanta University.

I have received permission to conduct my research in your school system. Enclosed is a copy of that confirmation.

I sincerely hope that you can spare a few minutes of your time and also that of your staff to aid me in obtaining information necessary for the completion of this study. I assure you that your school, selected teachers, and students will have complete anonymity.

If you object to my using your school in this investigation, please notify me at the above address within the next three days. If I do not hear from you, I will assume that I have your permission to conduct the study at your school.

Thank you in advance for your time and consideration.

Sincerely yours,

Mrs. Anne L. Crawford

Enclosure

Mrs. Anne L. Crawford
1423 Mill Acres Drive, S. W.
Atlanta, Georgia 30311

Dear Sir:

In order to complete requirements for the Ed.D. degree at Atlanta University, I am conducting a research study on Teachers' Perceptions of the Goal-Setting Model and the Traditional Model of Teacher Evaluation in Conjuction with Student Achievement; however, in order to gather data for my dissertation, I need your permission to conduct an experimental study in two middle schools in Area III.

If permission in granted, I will randomly select two middle schools for this study, and teachers will be asked to respond to the Teacher Questionnaire and administer the California Achievement Test. I assure you that the names of the selected schools, teachers and students will remain anonymous.

Please complete the enclosed response form and mail it to me at the above address as soon as possible.

Thank you for your cooperation.

Sincerely yours,

Mrs. Anne L. Crawford

Enclosure

RESPONSE FORM

TO: Anne Crawford
1423 Mill Acres Drive, S. W.
Atlanta, Georgia 30311

FROM: _____ School District

SUBJECT: Teachers' Perceptions of the Goal-Setting and the Traditional Models of Teacher Evaluation in Conjunction with Student Achievement

DATE: _____, 1985

_____ You have my permission to use two (2) randomly selected middle schools in Area III in the Atlanta Public School System in your research study if the principals agree with your request.

_____ Permission is not granted to use two (2) middle schools in the Atlanta Public School System in your research study.

SIGNED: _____

TITLE: _____

Traditional Model
(Procedure)

1. During preplanning week a copy of the evaluative instrument will be issued to all certified personnel. At that time, the philosophy upon which the instrument was constructed will be explained. Certified personnel will be evaluated each year.

   At the initial meeting, with all personnel being evaluated, the principal evaluator will explain the evaluative process and issue the necessary forms.

2. In September the evaluatee will collect data ib assigned students. The evaluatee will meet with the evaluator no later than the third week in September to reach a concensus on school system objectives, and a plan of action.

3. The first visitation/assessment will be held no later than the second week in October. This visitation/conference will determine the extent of which the evaluatee is implementing and realizing the establised objectives.

   a. If progress is satisfactory, the evaluatee proceeds with the implementation of the plans and consults with the evaluator, if needed.

   b. If progress is not satisfactory, the evaluator will assist the evaluatee.

4. During the month of November, the evaluatee will work on objectives and consult with the evaluator.

5. During the month of December, the evaluator will make his/her visit to the evaluatee's classroom, which will conclude with a conference between the evaluator and evaluatee. If progress is satisfactory, the evaluatee proceeds

with the implementation of plans, and if not consults with the evaluator, if needed.

6. In January, the evaluatee will work on objectives, consulting with the evaluator when necessary.

7. In February, the evaluator again observes the evaluatee in the classroom and consults and advises. Immediately after the third observation, the evaluator will request a conference with the evaluatee. At that time, the evaluator will write a special summative report on the evaluatee, including specific recommendations. Signed copies will be given to the evaluatee, evaluator, and the division of personnel.

8. All work sheets and observation forms will be printed in duplicate with the evaluator and evaluatee keeping copies on file.

# APPENDIX B

# INSTRUMENTS

# APPENDIX B

## INSTRUMENTS

### Teacher Evaulation Questionnaire

The purpose of this questionnaire is to assess, via your perceptions, the degree to which the existing planning and review process for the evaluation/ supervision of teachers is consistent with recognized and accepted practices and research in the area of evaluation/supervision of teachers as they are described in recent literature.

Below are listed criteria developed from a review of recent literature relating to the evaluation/supervision of teachers. Please indicate the degree to which you agree or disagree that the current process meets each criteria.

| | Strongly Agree | Agree | Disagree | Strongly Disagree |
|---|---|---|---|---|
| 1. Provides for teacher self appraisal | _____ | _____ | _____ | _____ |
| 2. Provides for stated expectations for teacher performance | _____ | _____ | _____ | _____ |
| 3. Provides for job goal-setting and attainment measurement | _____ | _____ | _____ | _____ |

| | Strongly Agree | Agree | Disagree | Strongly Agree |
|---|---|---|---|---|
| 4. Provides for evaluation in areas other than observed in classroom performance | _____ | _____ | _____ | _____ |
| 5. Provides for the evualuated to have input into the development/revision of the process | _____ | _____ | _____ | _____ |
| 6. Provides for systematic contact between teacher and supervisor | _____ | _____ | _____ | _____ |
| 7. Provides for the identification and reinforcement of teacher strengths | _____ | _____ | _____ | _____ |
| 8. Provides for the identification of problems such as substandard performance | _____ | _____ | _____ | _____ |
| 9. Provides for the relationship of teacher behavior to school district and building goals | _____ | _____ | _____ | _____ |

| | Strongly Agree | Agree | Disagree | Strongly Agree |
|---|---|---|---|---|
| 10. Provides for specific direction for the improvement of teacher behaviors | _____ | _____ | _____ | _____ |
| 11. Is specific in the mechanics of implementation of goals | _____ | _____ | _____ | _____ |
| 12. Provides/ Integrates individual performance objectives with the goals and objectives of the school organization | _____ | _____ | _____ | _____ |
| 13. Provides for higher student achievement | _____ | _____ | _____ | _____ |

EVALUATION SUMMARY SHEET FOR TRADITIONAL MODEL

EVALUATEE_____DATE _____

SCHOOL/DEPARTMENT _____POSITION _____

TENURE STATUS _____ PROBATIONARY STATUS:

1st Year _____2nd Year _____ 3rd _____

TEACHING PERFORMANCE EVALUATOR'S COMMENTS

I. Instructional Skills

II. Classroom Management and Organziation

VIII. Working Relationships

IV. Professional Growth and Responsibility

V. Personal Qualities

VI. Limiting Conditions

VII. Summary Statements Concerning Mutually Agreed Upon Objectives

Objectives

Plan of Action

Method of Assessment

RECOMMENDATIONS:

Number of Observation _____
Dates of Observation _____
Evaluatee is Recommended for Retention? Yes __ No __

Signature of Evaluator _____ Date _____

Signature of Evaluatee _____ Date _____

Signature of Reviewer _____ Date _____